

DEX0477US.WP

AK

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
30 January 2003 (30.01.2003)

PCT

(10) International Publication Number  
WO 2003/008537 A2

(51) International Patent Classification<sup>7</sup>: C12N Hills, CA 91367 (US). XIE, Zhidong; 22228 Victory Blvd., H-111, Woodland Hills, CA 91367 (US).

(21) International Application Number: PCT/US2002/010189 (74) Agent: TAHMASSEBI, Sam; Knobbe, Martens, Olson & Bear, Llp, 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(22) International Filing Date: 29 March 2002 (29.03.2002)

(25) Filing Language: English (81) Designated State (national): AU.

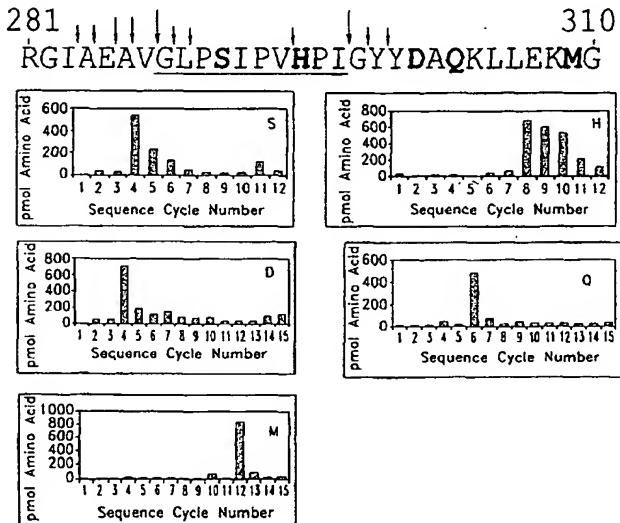
(26) Publication Language: English Published:  
— without international search report and to be republished upon receipt of that report

(30) Priority Data:  
60/282,211 6 April 2001 (06.04.2001) US  
60/337,017 7 November 2001 (07.11.2001) US  
60/363,210 7 March 2002 (07.03.2002) US (48) Date of publication of this corrected version:  
19 February 2004

(71) Applicant: MANNKIND CORPORATION [US/US];  
28903 North Avenue Paine, Valencia, CA 91355 (US). (15) Information about Correction:  
see PCT Gazette No. 08/2004 of 19 February 2004, Section II

(72) Inventors: SIMARD, John, J.L.; 11918 Laughton Way,  
Northridge, CA 91326 (US). DIAMOND, David, C.;  
23135 Schoenborn Street, West Hills, CA 91304 (US).  
LIU, Liping; 22228 Victory Blvd., H-111, Woodland (49) Date of filing of this corrected version:  
19 February 2004

(54) Title: EPITOPE SEQUENCES



Pool sequencing of PSMA\_281\_310 Digested for 60 min by Proteasome

(57) Abstract: Disclosed herein are polypeptides, including epitopes, clusters, and antigens. Also disclosed are compositions including said polypeptides and methods for their use.

BEST AVAILABLE COPY

WO 2003/008537 A2

## EPITOPE SEQUENCES

Field of the Invention

5 The present invention generally relates to peptides, and nucleic acids encoding peptides, that are useful epitopes of target-associated antigens. More specifically, the invention relates to epitopes that have a high affinity for MHC class I and that are produced by target-specific proteasomes.

Description of the Related ArtNeoplasia and the Immune System

10 The neoplastic disease state commonly known as cancer is thought to result generally from a single cell growing out of control. The uncontrolled growth state typically results from a multi-step process in which a series of cellular systems fail, resulting in the genesis of a neoplastic cell. The resulting neoplastic cell rapidly reproduces itself, forms one or more tumors, and eventually may cause the death of the host.

15 Because the progenitor of the neoplastic cell shares the host's genetic material, neoplastic cells are largely unassailed by the host's immune system. During immune surveillance, the process in which the host's immune system surveys and localizes foreign materials, a neoplastic cell will appear to the host's immune surveillance machinery as a "self" cell.

Viruses and the Immune System

20 In contrast to cancer cells, virus infection involves the expression of clearly non-self antigens. As a result, many virus infections are successfully dealt with by the immune system with minimal clinical sequelae. Moreover, it has been possible to develop effective vaccines for many of those infections that do cause serious disease. A variety of vaccine approaches have been used successfully to combat various diseases. These approaches include subunit vaccines consisting of individual proteins produced through recombinant DNA technology. Notwithstanding these 25 advances, the selection and effective administration of minimal epitopes for use as viral vaccines has remained problematic.

30 In addition to the difficulties involved in epitope selection stands the problem of viruses that have evolved the capability of evading a host's immune system. Many viruses, especially viruses that establish persistent infections, such as members of the herpes and pox virus families, produce immunomodulatory molecules that permit the virus to evade the host's immune system. The effects of these immunomodulatory molecules on antigen presentation may be overcome by the targeting of select epitopes for administration as immunogenic compositions. To better understand the interaction of neoplastic cells and virally infected cells with the host's immune system, a discussion of the system's components follows below.

35 The immune system functions to discriminate molecules endogenous to an organism ("self" molecules) from material exogenous or foreign to the organism ("non-self" molecules). The

immune system has two types of adaptive responses to foreign bodies based on the components that mediate the response: a humoral response and a cell-mediated response. The humoral response is mediated by antibodies, while the cell-mediated response involves cells classified as lymphocytes. Recent anticancer and antiviral strategies have focused on mobilizing the host immune system as a means of anticancer or antiviral treatment or therapy.

5

The immune system functions in three phases to protect the host from foreign bodies: the cognitive phase, the activation phase, and the effector phase. In the cognitive phase, the immune system recognizes and signals the presence of a foreign antigen or invader in the body. The foreign antigen can be, for example, a cell surface marker from a neoplastic cell or a viral protein. Once the system is aware of an invading body, antigen specific cells of the immune system proliferate and differentiate in response to the invader-triggered signals. The last stage is the effector stage in which the effector cells of the immune system respond to and neutralize the detected invader.

10

An array of effector cells implements an immune response to an invader. One type of effector cell, the B cell, generates antibodies targeted against foreign antigens encountered by the host. In combination with the complement system, antibodies direct the destruction of cells or organisms bearing the targeted antigen. Another type of effector cell is the natural killer cell (NK cell), a type of lymphocyte having the capacity to spontaneously recognize and destroy a variety of virus infected cells as well as malignant cell types. The method used by NK cells to recognize target cells is poorly understood.

15

Another type of effector cell, the T cell, has members classified into three subcategories, each playing a different role in the immune response. Helper T cells secrete cytokines which stimulate the proliferation of other cells necessary for mounting an effective immune response, while suppressor T cells down-regulate the immune response. A third category of T cell, the cytotoxic T cell (CTL), is capable of directly lysing a targeted cell presenting a foreign antigen on its surface.

20

#### The Major Histocompatibility Complex and T Cell Target Recognition

25

T cells are antigen-specific immune cells that function in response to specific antigen signals. B lymphocytes and the antibodies they produce are also antigen-specific entities. However, unlike B lymphocytes, T cells do not respond to antigens in a free or soluble form. For a T cell to respond to an antigen, it requires the antigen to be processed to peptides which are then bound to a presenting structure encoded in the major histocompatibility complex (MHC). This requirement is called "MHC restriction" and it is the mechanism by which T cells differentiate "self" from "non-self" cells. If an antigen is not displayed by a recognizable MHC molecule, the T cell will not recognize and act on the antigen signal. T cells specific for a peptide bound to a recognizable MHC molecule bind to these MHC-peptide complexes and proceed to the next stages of the immune response.

30

There are two types of MHC, class I MHC and class II MHC. T Helper cells (CD4<sup>+</sup>) predominately interact with class II MHC proteins while cytolytic T cells (CD8<sup>+</sup>) predominately

interact with class I MHC proteins. Both classes of MHC protein are transmembrane proteins with a majority of their structure on the external surface of the cell. Additionally, both classes of MHC proteins have a peptide binding cleft on their external portions. It is in this cleft that small fragments of proteins, endogenous or foreign, are bound and presented to the extracellular environment.

5 Cells called "professional antigen presenting cells" (pAPCs) display antigens to T cells using the MHC proteins but additionally express various co-stimulatory molecules depending on the particular state of differentiation/activation of the pAPC. When T cells, specific for the peptide bound to a recognizable MHC protein, bind to these MHC-peptide complexes on pAPCs, the specific co-stimulatory molecules that act upon the T cell direct the path of differentiation/activation taken by the  
10 T cell. That is, the co-stimulation molecules affect how the T cell will act on antigenic signals in future encounters as it proceeds to the next stages of the immune response.

15 As discussed above, neoplastic cells are largely ignored by the immune system. A great deal of effort is now being expended in an attempt to harness a host's immune system to aid in combating the presence of neoplastic cells in a host. One such area of research involves the formulation of anticancer vaccines.

#### Anticancer Vaccines

20 Among the various weapons available to an oncologist in the battle against cancer is the immune system of the patient. Work has been done in various attempts to cause the immune system to combat cancer or neoplastic diseases. Unfortunately, the results to date have been largely disappointing. One area of particular interest involves the generation and use of anticancer vaccines.

25 To generate a vaccine or other immunogenic composition, it is necessary to introduce to a subject an antigen or epitope against which an immune response may be mounted. Although neoplastic cells are derived from and therefore are substantially identical to normal cells on a genetic level, many neoplastic cells are known to present tumor-associated antigens (TuAAs). In theory, these antigens could be used by a subject's immune system to recognize these antigens and attack the neoplastic cells. In reality, however, neoplastic cells generally appear to be ignored by the host's immune system.

30 A number of different strategies have been developed in an attempt to generate vaccines with activity against neoplastic cells. These strategies include the use of tumor-associated antigens as immunogens. For example, U.S. Patent No. 5,993,828, describes a method for producing an immune response against a particular subunit of the Urinary Tumor Associated Antigen by administering to a subject an effective dose of a composition comprising inactivated tumor cells having the Urinary Tumor Associated Antigen on the cell surface and at least one tumor associated antigen selected from the group consisting of GM-2, GD-2, Fetal Antigen and Melanoma  
35

Associated Antigen. Accordingly, this patent describes using whole, inactivated tumor cells as the immunogen in an anticancer vaccine.

Another strategy used with anticancer vaccines involves administering a composition containing isolated tumor antigens. In one approach, MAGE-A1 antigenic peptides were used as an immunogen. (See Chaux, P., *et al.*, "Identification of Five MAGE-A1 Epitopes Recognized by Cytolytic T Lymphocytes Obtained by *In Vitro* Stimulation with Dendritic Cells Transduced with MAGE-A1," *J. Immunol.*, 163(5):2928-2936 (1999)). There have been several therapeutic trials using MAGE-A1 peptides for vaccination, although the effectiveness of the vaccination regimes was limited. The results of some of these trials are discussed in Vose, J.M., "Tumor Antigens Recognized by T Lymphocytes," 10<sup>th</sup> European Cancer Conference, Day 2, Sept. 14, 1999.

In another example of tumor associated antigens used as vaccines, Scheinberg, *et al.* treated 12 chronic myelogenous leukemia (CML) patients already receiving interferon (IFN) or hydroxyurea with 5 injections of class I-associated bcr-abl peptides with a helper peptide plus the adjuvant QS-21. Scheinberg, D.A., *et al.*, "BCR-ABL Breakpoint Derived Oncogene Fusion Peptide Vaccines Generate Specific Immune Responses in Patients with Chronic Myelogenous Leukemia (CML) [Abstract 1665], American Society of Clinical Oncology 35<sup>th</sup> Annual Meeting, Atlanta (1999). Proliferative and delayed type hypersensitivity (DTH) T cell responses indicative of T-helper activity were elicited, but no cytolytic killer T cell activity was observed within the fresh blood samples.

Additional examples of attempts to identify TuAAs for use as vaccines are seen in the recent work of Cebon, *et al.* and Scheibenbogen, *et al.* Cebon, *et al.* immunized patients with metastatic melanoma using intradermally administered MART-1<sub>26-35</sub> peptide with IL-12 in increasing doses given either subcutaneously or intravenously. Of the first 15 patients, 1 complete remission, 1 partial remission, and 1 mixed response were noted. Immunc assays for T cell generation included DTH, which was seen in patients with or without IL-12. Positive CTL assays were seen in patients with evidence of clinical benefit, but not in patients without tumor regression. Cebon, *et al.*, "Phase I Studies of Immunization with Melan-A and IL-12 in HLA A2+ Positive Patients with Stage III and IV Malignant Melanoma," [Abstract 1671], American Society of Clinical Oncology 35<sup>th</sup> Annual Meeting, Atlanta (1999).

Scheibenbogen, *et al.* immunized 18 patients with 4 HLA class I restricted tyrosinase peptides, 16 with metastatic melanoma and 2 adjuvant patients. Scheibenbogen, *et al.*, "Vaccination with Tyrosinase peptides and GM-CSF in Metastatic Melanoma: a Phase II Trial," [Abstract 1680], American Society of Clinical Oncology 35<sup>th</sup> Annual Meeting, Atlanta (1999). Increased CTL activity was observed in 4/15 patients, 2 adjuvant patients, and 2 patients with evidence of tumor regression. As in the trial by Cebon, *et al.*, patients with progressive disease did

not show boosted immunity. In spite of the various efforts expended to date to generate efficacious anticancer vaccines, no such composition has yet been developed.

#### Antiviral Vaccines

5 Vaccine strategies to protect against viral diseases have had many successes. Perhaps the most notable of these is the progress that has been made against the disease small pox, which has been driven to extinction. The success of the polio vaccine is of a similar magnitude.

10 Viral vaccines can be grouped into three classifications: live attenuated virus vaccines, such as vaccinia for small pox, the Sabin poliovirus vaccine, and measles mumps and rubella; whole killed or inactivated virus vaccines, such as the Salk poliovirus vaccine, hepatitis A virus vaccine and the typical influenza virus vaccines; and subunit vaccines, such as hepatitis B. Due to their lack of a complete viral genome, subunit vaccines offer a greater degree of safety than those based on whole viruses.

15 The paradigm of a successful subunit vaccine is the recombinant hepatitis B vaccine based on the viruses envelope protein. Despite much academic interest in pushing the reductionist subunit concept beyond single proteins to individual epitopes, the efforts have yet to bear much fruit. Viral vaccine research has also concentrated on the induction of an antibody response although cellular responses also occur. However, many of the subunit formulations are particularly poor at generating a CTL response.

#### Summary of the Invention

20 Previous methods of priming professional antigen presenting cells (pAPCs) to display target cell epitopes have relied simply on causing the pAPCs to express target-associated antigens (TAAs), or epitopes of those antigens which are thought to have a high affinity for MHC I molecules. However, the proteasomal processing of such antigens results in presentation of epitopes on the pAPC that do not correspond to the epitopes present on the target cells.

25 Using the knowledge that an effective cellular immune response requires that pAPCs present the same epitope that is presented by the target cells, the present invention provides epitopes that have a high affinity for MHC I, and that correspond to the processing specificity of the housekeeping proteasome, which is active in peripheral cells. These epitopes thus correspond to those presented on target cells. The use of such epitopes in vaccines can activate the cellular immune response to recognize the correctly processed TAA and can result in removal of target cells that present such epitopes. In some embodiments, the housekeeping epitopes provided herein can be used in combination with immune epitopes, generating a cellular immune response that is competent to attack target cells both before and after interferon induction. In other embodiments the epitopes are useful in the diagnosis and monitoring of the target-associated disease and in the 30 35 generation of immunological reagents for such purposes.

5 Embodiments of the invention relate to isolated epitopes and antigens or polypeptides that comprise the epitopes. Preferred embodiments include an epitope or antigen having the sequence as disclosed in TABLE 1. Other embodiments can include an epitope cluster comprising a polypeptide from Table 1. Further, embodiments include a polypeptide having substantial similarity to the already mentioned epitopes, antigens, or clusters. Other preferred embodiments include a polypeptide having functional similarity to any of the above. Still further embodiments relate to a nucleic acid encoding the polypeptide of any of the epitopes, clusters, antigens, and polypeptides from Table 1 and mentioned herein.

10 The epitope can be immunologically active. The polypeptide comprising the epitope can be less than about 30 amino acids in length, more preferably, the polypeptide is 8 to 10 amino acids in length, for example. Substantial or functional similarity can include addition of at least one amino acid, for example, and the at least one additional amino acid can be at an N-terminus of the polypeptide. The substantial or functional similarity can include a substitution of at least one amino acid.

15 The epitope, cluster, or polypeptide comprising the same can have affinity to an HLA-A2 molecule. The affinity can be determined by an assay of binding, by an assay of restriction of epitope recognition, by a prediction algorithm, and the like. The epitope, cluster, or polypeptide comprising the same can have affinity to an HLA-B7, HLA-B51 molecule, and the like.

20 In preferred embodiments the polypeptide can be a housekeeping epitope. The epitope or polypeptide can correspond to an epitope displayed on a tumor cell, to an epitope displayed on a neovasculature cell, and the like. The epitope or polypeptide can be an immune epitope. The epitope, cluster and/or polypeptide can be a nucleic acid.

25 Other embodiments relate to pharmaceutical compositions comprising the polypeptides, including an epitope from Table 1, a cluster, or a polypeptide comprising the same and a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like. The adjuvant can be a polynucleotide. The polynucleotide can include a dinucleotide. The dinucleotide can be CpG, for example. The adjuvant can be encoded by a polynucleotide. The adjuvant can be a cytokine and the cytokine can be, for example, GM-CSF.

30 The pharmaceutical compositions can further include a professional antigen-presenting cell (pAPC). The pAPC can be a dendritic cell, for example. The pharmaceutical composition can further include a second epitope. The second epitope can be a polypeptide. The second epitope can be a nucleic acid. The second epitope can be a housekeeping epitope, an immune epitope, and the like.

35 Still further embodiments relate to pharmaceutical compositions that include any of the nucleic acids discussed herein, including those that encode polypeptides that comprise epitopes or

antigens from Table 1. Such compositions can include a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like.

5 Other embodiments relate to recombinant constructs that include such a nucleic acid as described herein, including those that encode polypeptides that comprise epitopes or antigens from Table 1. The constructs can further include a plasmid, a viral vector, an artificial chromosome, and the like. The construct can further include a sequence encoding at least one feature, such as for example, a second epitope, an IRES, an ISS, an NIS, ubiquitin.

10 Further embodiments relate to purified antibodies that specifically bind to at least one of the epitopes in Table 1A. Other embodiments relate to purified antibodies that specifically bind to a peptide-MHC protein complex comprising an epitope disclosed in Table 1A or any other suitable epitope. The antibody from any embodiment can be a monoclonal antibody.

Still other embodiments relate to multimeric MHC-peptide complexes that include an epitope, such as, for example, an epitope disclosed in Table 1.

15 Embodiments relate to isolated T cells expressing a T cell receptor specific for an MHC-peptide complex. The complex can include an epitope, such as, for example, an epitope disclosed in Table 1 of claim 1. The T cell can be produced by an *in vitro* immunization. The T cell can be isolated from an immunized animal. Embodiments relate to T cell clones, including cloned T cells, such as those discussed above. Embodiments also relate to polyclonal population of T cells. Such populations can include a T cell, as described above, for example.

20 Still further embodiments relate to pharmaceutical compositions that include a T cell, such as those described above, for example, and a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like.

25 Embodiments of the invention relate to isolated protein molecules comprising the binding domain of a T cell receptor specific for an MHC-peptide complex. The complex can include an epitope disclosed in Table 1. The protein can be multivalent. Other embodiments relate to isolated nucleic acids encoding such proteins. Still further embodiments relate to recombinant constructs that include such nucleic acids.

30 Other embodiments of the invention relate to host cells expressing the recombinant construct described herein, including constructs encoding an epitope, cluster or polypeptide comprising the same, disclosed in Table 1, for example. The host cell can be a dendritic cell, macrophage, tumor cell, tumor-derived cell, and the like. The host cell can be a bacterium, fungus, protozoan and the like. Embodiments also relate to pharmaceutical compositions that include a host cell, such as those discussed herein, and a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like.

35 Still other embodiments relate to vaccines or immunotherapeutic compositions that include at least one component, such as, for example, an epitope disclosed in Table 1 or otherwise

described herein; a cluster that includes such an epitope, an antigen or polypeptide that includes such an epitope; a composition described above and herein; a construct, a T cell, or a host cell as described above and herein.

Further embodiments relate to methods of treating an animal. The methods can include 5 administering to an animal a vaccine or immunotherapeutic composition, including those disclosed above and herein. The administering step can include a mode of delivery, such as, for example, transdermal, intranodal, perinodal, oral, intravenous, intradermal, intramuscular, intraperitoneal, mucosal, aerosol inhalation, instillation, and the like. The method can further include a step of assaying to determine a characteristic indicative of a state of a target cell or target cells. The 10 method can include a first assaying step and a second assaying step, wherein the first assaying step precedes the administering step, and wherein the second assaying step follows the administering step. The method can further include a step of comparing the characteristic determined in the first assaying step with the characteristic determined in the second assaying step to obtain a result. The 15 result can be for example, evidence of an immune response, a diminution in number of target cells, a loss of mass or size of a tumor comprising target cells, a decrease in number or concentration of an intracellular parasite infecting target cells, and the like.

Embodiments relate to methods of evaluating immunogenicity of a vaccine or immunotherapeutic composition. The methods can include administering to an animal a vaccine or immunotherapeutic, such as those described above and elsewhere herein, and evaluating 20 immunogenicity based on a characteristic of the animal. The animal can be HLA-transgenic.

Other embodiments relate to methods of evaluating immunogenicity that include *in vitro* stimulation of a T cell with the vaccine or immunotherapeutic composition, such as those described above and elsewhere herein, and evaluating immunogenicity based on a characteristic of the T cell. The stimulation can be a primary stimulation.

25 Still further embodiments relate to methods of making a passive/adoptive immunotherapeutic. The methods can include combining a T cell or a host cell, such as those described above and elsewhere herein, with a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like.

Other embodiments relate to methods of determining specific T cell frequency, and can 30 include the step of contacting T cells with a MHC-peptide complex comprising an epitope disclosed in Table 1, or a complex comprising a cluster or antigen comprising such an epitope. The contacting step can include at least one feature, such as, for example, immunization, restimulation, detection, enumeration, and the like. The method can further include ELISPOT analysis, limiting dilution analysis, flow cytometry, *in situ* hybridization, the polymerase chain reaction, any 35 combination thereof, and the like.

Embodiments relate to methods of evaluating immunologic response. The methods can include the above-described methods determining specific T cell frequency carried out prior to and subsequent to an immunization step.

5 Another embodiment relates to methods of evaluating immunologic response. The methods can include determining frequency, cytokine production, or cytolytic activity of T cells, prior to and subsequent to a step of stimulation with MHC-peptide complexes comprising an epitope, such as, for example an epitope from Table 1, a cluster or a polypeptide comprising such an epitope.

10 Further embodiments relate to methods of diagnosing a disease. The methods can include contacting a subject tissue with at least one component, including, for example, a T cell, a host cell, an antibody, a protein, including those described above and elsewhere herein; and diagnosing the disease based on a characteristic of the tissue or of the component. The contacting step can take place *in vivo*. The contacting step can take place *in vitro*.

15 Still other embodiments relate to methods of making a vaccine. The methods can include combining at least one component, an epitope, a composition, a construct, a T cell, a host cell; including any of those described above and elsewhere herein, with a pharmaceutically acceptable adjuvant, carrier, diluent, excipient, and the like.

20 Embodiments relate to computer readable media having recorded thereon the sequence of any one of SEQ ID NOS: 1 -602, in a machine having a hardware or software that calculates the physical, biochemical, immunologic, or molecular genetic properties of a molecule embodying said sequence.

25 Still other embodiments relate to methods of treating an animal. The methods can include combining the method of treating an animal that includes administering to the animal a vaccine or immunotherapeutic composition, such as described above and elsewhere herein, combined with at least one mode of treatment, including, for example, radiation therapy, chemotherapy, biochemotherapy, surgery, and the like.

30 Further embodiments relate to isolated polypeptides that include an epitope cluster from a target-associated antigen having the sequence as disclosed in any one of Tables 25-44, wherein the amino acid sequence includes not more than about 80% of the amino acid sequence of the antigen.

35 Other embodiments relate to vaccines or immunotherapeutic products that include an isolated peptide as described above and elsewhere herein. Still other embodiments relate to isolated polynucleotides encoding a polypeptide as described above and elsewhere herein. Other embodiments relate vaccines or immunotherapeutic products that include these polynucleotides. The polynucleotide can be DNA or RNA.

Brief Description of the Drawings

Figure 1 is a sequence alignment of NY-ESO-1 and several similar protein sequences.

Figure 2 graphically represents a plasmid vaccine backbone useful for delivering nucleic acid-encoded epitopes.

5 Figures 3A and 3B are FACS profiles showing results of HLA-A2 binding assays for tyrosinase<sub>207-215</sub> and tyrosinase<sub>208-216</sub>.

Figure 4 is a T=120 min. time point mass spectrum of the fragments produced by proteasomal cleavage of SSX-2<sub>31-68</sub>.

Figure 5 shows a binding curve for HLA-A2:SSX-2<sub>41-49</sub> with controls.

10 Figure 6 shows specific lysis of SSX-2<sub>41-49</sub>-pulsed targets by CTL from SSX-2<sub>41-49</sub>-immunized HLA-A2 transgenic mice.

Figure 7A, B, and C show results of N-terminal pool sequencing of a T=60 min. time point aliquot of the PSMA<sub>163-192</sub> proteasomal digest.

15 Figure 8 shows binding curves for HLA-A2:PSMA<sub>168-177</sub> and HLA-A2:PSMA<sub>288-297</sub> with controls.

Figure 9 shows results of N-terminal pool sequencing of a T=60 min. time point aliquot of the PSMA<sub>281-310</sub> proteasomal digest.

Figure 10 shows binding curves for HLA-A2:PSMA<sub>461-469</sub>, HLA-A2:PSMA<sub>460-469</sub>, and HLA-A2:PSMA<sub>663-671</sub>, with controls.

20 Figure 11 shows the results of a  $\gamma$ -IFN-based ELISPOT assay detecting PSMA<sub>463-471</sub>-reactive HLA-A1<sup>+</sup> CD8<sup>+</sup> T cells.

Figure 12 shows blocking of reactivity of the T cells used in figure 10 by anti-HLA-A1 mAb, demonstrating HLA-A1-restricted recognition.

Figure 13 shows a binding curve for HLA-A2:PSMA<sub>663-671</sub>, with controls.

25 Figure 14 shows a binding curve for HLA-A2:PSMA<sub>662-671</sub>, with controls.

Figure 15. Comparison of anti-peptide CTL responses following immunization with various doses of DNA by different routes of injection.

Figure 16. Growth of transplanted gp33 expressing tumor in mice immunized by i.ln. injection of gp33 epitope-expressing, or control, plasmid.

30 Figure 17. Amount of plasmid DNA detected by real-time PCR in injected or draining lymph nodes at various times after i.ln. of i.m. injection, respectively.

Detailed Description of the Preferred EmbodimentDefinitions

Unless otherwise clear from the context of the use of a term herein, the following listed 35 terms shall generally have the indicated meanings for purposes of this description.

PROFESSIONAL ANTIGEN-PRESENTING CELL (pAPC) – a cell that possesses T cell costimulatory molecules and is able to induce a T cell response. Well characterized pAPCs include dendritic cells, B cells, and macrophages.

PERIPHERAL CELL – a cell that is not a pAPC.

5 HOUSEKEEPING PROTEASOME – a proteasome normally active in peripheral cells, and generally not present or not strongly active in pAPCs.

IMMUNE PROTEASOME – a proteasome normally active in pAPCs; the immune proteasome is also active in some peripheral cells in infected tissues.

10 EPITOPE – a molecule or substance capable of stimulating an immune response. In preferred embodiments, epitopes according to this definition include but are not necessarily limited to a polypeptide and a nucleic acid encoding a polypeptide, wherein the polypeptide is capable of stimulating an immune response. In other preferred embodiments, epitopes according to this definition include but are not necessarily limited to peptides presented on the surface of cells, the peptides being non-covalently bound to the binding cleft of class I MHC, such that they can 15 interact with T cell receptors.

MHC EPITOPE – a polypeptide having a known or predicted binding affinity for a mammalian class I or class II major histocompatibility complex (MHC) molecule.

20 HOUSEKEEPING EPITOPE – In a preferred embodiment, a housekeeping epitope is defined as a polypeptide fragment that is an MHC epitope, and that is displayed on a cell in which housekeeping proteasomes are predominantly active. In another preferred embodiment, a housekeeping epitope is defined as a polypeptide containing a housekeeping epitope according to the foregoing definition, that is flanked by one to several additional amino acids. In another preferred embodiment, a housekeeping epitope is defined as a nucleic acid that encodes a housekeeping epitope according to the foregoing definitions.

25 IMMUNE EPITOPE – In a preferred embodiment, an immune epitope is defined as a polypeptide fragment that is an MHC epitope, and that is displayed on a cell in which immune proteasomes are predominantly active. In another preferred embodiment, an immune epitope is defined as a polypeptide containing an immune epitope according to the foregoing definition, that is flanked by one to several additional amino acids. In another preferred embodiment, an immune epitope is defined as a polypeptide including an epitope cluster sequence, having at least two polypeptide sequences having a known or predicted affinity for a class I MHC. In yet another preferred embodiment, an immune epitope is defined as a nucleic acid that encodes an immune epitope according to any of the foregoing definitions.

30 TARGET CELL – a cell to be targeted by the vaccines and methods of the invention. Examples of target cells according to this definition include but are not necessarily limited to: a

neoplastic cell and a cell harboring an intracellular parasite, such as, for example, a virus, a bacterium, or a protozoan.

5 TARGET-ASSOCIATED ANTIGEN (TAA) – a protein or polypeptide present in a target cell.

5 TUMOR-ASSOCIATED ANTIGENS (TuAA) – a TAA, wherein the target cell is a neoplastic cell.

HLA EPITOPE – a polypeptide having a known or predicted binding affinity for a human class I or class II HLA complex molecule.

10 ANTIBODY – a natural immunoglobulin (Ig), poly- or monoclonal, or any molecule composed in whole or in part of an Ig binding domain, whether derived biochemically or by use of recombinant DNA. Examples include *inter alia*, F(ab), single chain Fv, and Ig variable region-phage coat protein fusions.

15 ENCODE – an open-ended term such that a nucleic acid encoding a particular amino acid sequence can consist of codons specifying that (poly)peptide, but can also comprise additional sequences either translatable, or for the control of transcription, translation, or replication, or to facilitate manipulation of some host nucleic acid construct.

20 SUBSTANTIAL SIMILARITY – this term is used to refer to sequences that differ from a reference sequence in an inconsequential way as judged by examination of the sequence. Nucleic acid sequences encoding the same amino acid sequence are substantially similar despite differences in degenerate positions or modest differences in length or composition of any non-coding regions. Amino acid sequences differing only by conservative substitution or minor length variations are substantially similar. Additionally, amino acid sequences comprising housekeeping epitopes that differ in the number of N-terminal flanking residues, or immune epitopes and epitope clusters that differ in the number of flanking residues at either terminus, are substantially similar. Nucleic acids that encode substantially similar amino acid sequences are themselves also substantially similar.

25 FUNCTIONAL SIMILARITY – this term is used to refer to sequences that differ from a reference sequence in an inconsequential way as judged by examination of a biological or biochemical property, although the sequences may not be substantially similar. For example, two nucleic acids can be useful as hybridization probes for the same sequence but encode differing amino acid sequences. Two peptides that induce cross-reactive CTL responses are functionally similar even if they differ by non-conservative amino acid substitutions (and thus do not meet the substantial similarity definition). Pairs of antibodies, or TCRs, that recognize the same epitope can be functionally similar to each other despite whatever structural differences exist. In testing for functional similarity of immunogenicity one would generally immunize with the “altered” antigen and test the ability of the elicited response (Ab, CTL, cytokine production, etc.) to recognize the target antigen. Accordingly, two sequences may be designed to differ in certain respects while

retaining the same function. Such designed sequence variants are among the embodiments of the present invention.

**Table 1A. SEQ ID NOS.\* including epitopes in Examples 1-7, 13.**

SEQ ID NO	IDENTITY	SEQUENCE
1	Tyr 207-216	FLPWHRLFLL
2	Tyrosinase protein	Accession number**: P14679
3	SSX-2 protein	Accession number: NP_003138
4	PSMA protein	Accession number: NP_004467
5	Tyrosinase cDNA	Accession number: NM_000372
6	SSX-2 cDNA	Accession number: NM_003147
7	PSMA cDNA	Accession number: NM_004476
8	Tyr 207-215	FLPWHRLFL
9	Tyr 208-216	LPWHRLFLL
10	SSX-2 31-68	YFSKEEWEKMKASEKIFYVYVMKRKYEAMTKLGFK ATLP
11	SSX-2 32-40	FSKEEWEKM
12	SSX-2 39-47	KMKASEKIF
13	SSX-2 40-48	MKASEKIFY
14	SSX-2 39-48	KMKASEKIFY
15	SSX-2 41-49	KASEKIFYV
16	SSX-2 40-49	MKASEKIFYV
17	SSX-2 41-50	KASEKIFYVV
18	SSX-2 42-49	ASEKIFYVY
19	SSX-2 53-61	RKYEAMTKL
20	SSX-2 52-61	KRKYEAMTKL
21	SSX-2 54-63	KYEAMTKLG
22	SSX-2 55-63	YEAMTKLG
23	SSX-2 56-63	EAMTKLG
24	HBV18-27	FLPSDYFPSV
25	HLA-B44 binder	AEMGKYSFY
26	SSX-1 41-49	KYSEKISYV
27	SSX-3 41-49	KVSEKIVYV
28	SSX-4 41-49	KSSEKIVYV
29	SSX-5 41-49	KASEKIIYV
30	PSMA163-192	AFSPQGMPEGDLVYVNYARTEDFFKLERDM
31	PSMA 168-190	GMPEGDLVYVNYARTEDFFKLER
32	PSMA 169-177	MPEGDLVYV
33	PSMA 168-177	GMPEGDLVYV
34	PSMA 168-176	GMPEGDLVY
35	PSMA 167-176	QGMPEGDLVY
36	PSMA 169-176	MPEGDLVY

37	PSMA 171-179	EGDLVYVNY
38	PSMA 170-179	PEGDLVYVNY
39	PSMA 174-183	LVYVNYARTE
40	PSMA 177-185	VNYARTEDF
41	PSMA 176-185	YVNYARTEDF
42	PSMA 178-186	NYARTEDFF
43	PSMA 179-186	YARTEDFF
44	PSMA 181-189	RTEDFFKLE
45	PSMA 281-310	RGIAEAVGLPSIPVHPIGYYYDAQKLLEKMG
46	PSMA 283-307	IAEAVGIPVHPIGYYYDAQKLLE
47	PSMA 289-297	LPSIPVHPI
48	PSMA 288-297	GLPSIPVHPI
49	PSMA 297-305	IGYYDAQKL
50	PSMA 296-305	PIGYYDAQKL
51	PSMA 291-299	SIPVHPIGY
52	PSMA 290-299	PSIPVHPIGY
53	PSMA 292-299	IPVHPIGY
54	PSMA 299-307	YYDAQKLLE
55	PSMA454-481	SSIEGNYTLRVDTPLMYSVLVHLTKEL
56	PSMA 456-464	IEGNYTLRV
57	PSMA 455-464	SIEGNYTLRV
58	PSMA 457-464	EGNYTLRV
59	PSMA 461-469	TLRVDCTPL
60	PSMA 460-469	YTLRVDCTPL
61	PSMA 462-470	LRVDCTPLM
62	PSMA 463-471	RVDCTPLMY
63	PSMA 462-471	LRVDCTPLMY
64	PSMA653-687	FDKSNPIVLRMMNDQLMFLERAFIDPLGLPDRPFY
65	PSMA 660-681	VLRMMNDQLMFLERAFIDPLGL
66	PSMA 663-671	MMNDQLMFL
67	PSMA 662-671	RMMNDQLMFL
68	PSMA 662-670	RMMNDQLMF
69	Tyr 1-17	MLLAVLYCLLWSFQTSA

Table 1B. SEQ ID NOS.\* including epitopes in Examples 14 and 15.

SEQ ID NO	IDENTITY	SEQUENCE
70	GP100 protein	**Accession number: P40967
71	MAGE-1 protein	Accession number: P43355
72	MAGE-2 protein	Accession number: P43356
73	MAGE-3 protein	Accession number: P43357
74	NY-ESO-1 protein	Accession number: P78358
75	LAEG-1a protein	Accession number: CAA11116

76	LAGE-1b protein	Accession number: CAA11117
77	PRAME protein	Accession number: NP 006106
78	PSA protein	Accession number: P07288
79	PSCA protein	Accession number: O43653
80	GP100 cds	Accession number: U20093
81	MAGE-1 cds	Accession number: M77481
82	MAGE-2 cds	Accession number: L18920
83	MAGE-3 cds	Accession number: U03735
84	NY-ESO-1 cDNA	Accession number: U87459
85	PRAME cDNA	Accession number: NM 006115
86	PSA cDNA	Accession number: NM_001648
87	PSCA cDNA	Accession number: AF043498
88	GP100 630-638	LPHSSSHWL
89	GP100 629-638	QLPHSSSHWL
90	GP100 614-622	LIYRRRLMK
91	GP100 613-622	SLIYRRRLMK
92	GP100 615-622	IYRRRLMK
93	GP100 630-638	LPHSSSHWL
94	GP100 629-638	QLPHSSSHWL
95	MAGE-1 95-102	ESLFRAVI
96	MAGE-1 93-102	IILESLFRAVI
97	MAGE-1 93-101	IILESLFRAV
98	MAGE-1 92-101	CILESLFRAV
99	MAGE-1 92-100	CILESLFRA
100	MAGE-1 263-271	EFLWGPRAL
101	MAGE-1 264-271	FLWGPRAL
102	MAGE-1 264-273	FLWGPRALAE
103	MAGE-1 265-274	LWGPRALAAET
104	MAGE-1 268-276	PRALAETSY
105	MAGE-1 267-276	GPRALAETSY
106	MAGE-1 269-277	RALAETSYV
107	MAGE-1 271-279	LAETSYVKV
108	MAGE-1 270-279	ALAETSYVKV
109	MAGE-1 272-280	AETSYVKVL
110	MAGE-1 271-280	LAETSYVKVL
111	MAGE-1 274-282	TSYVKVLEY
112	MAGE-1 273-282	ETSYVKVLEY
113	MAGE-1 278-286	KVLEYVIKV
114	MAGE-1 168-177	SYVLVTCLGL
115	MAGE-1 169-177	YVLVTCLGL
116	MAGE-1 170-177	VLVTCLGL
117	MAGE-1 240-248	TQDLVQEKY

118	MAGE-1 239-248	LTQDLVQEKY
119	MAGE-1 232-240	YGEPRKLLT
120	MAGE-1 243-251	LVQEKYLEY
121	MAGE-1 242-251	DLVQEKYLEY
122	MAGE-1 230-238	SAYGEPRKL
123	MAGE-1 278-286	KVLEYVIKV
124	MAGE-1 277-286	VKVLEYVIKV
125	MAGE-1 276-284	YVKVLEYVI
126	MAGE-1 274-282	TSYVKVLEY
127	MAGE-1 273-282	ETSYVKVLEY
128	MAGE-1 283-291	VIKVSARVR
129	MAGE-1 282-291	YVIKVSARVR
130	MAGE-2 115-122	ELVHFLLL
131	MAGE-2 113-122	MVELVHFLLL
132	MAGE-2 109-116	ISRKMVEL
133	MAGE-2 108-116	AISRKMVEL
134	MAGE-2 107-116	AAISRKMVEL
135	MAGE-2 112-120	KMVELVHFL
136	MAGE-2 109-117	ISRKMVELV
137	MAGE-2 108-117	AISRKMVELV
138	MAGE-2 116-124	LVHFLLLKY
139	MAGE-2 115-124	ELVHFLLLKY
140	MAGE-2 111-119	RKMVELVHF
141	MAGE-2 158-166	LQLVFGIEV
142	MAGE-2 157-166	YLQLVFGIEV
143	MAGE-2 159-167	QLVFGIEVV
144	MAGE-2 158-167	LQLVFGIEVV
145	MAGE-2 164-172	IEVVEVVPI
146	MAGE-2 163-172	GIEVVEVVPI
147	MAGE-2 162-170	FGIEVVEVV
148	MAGE-2 154-162	ASEYLQLVF
149	MAGE-2 153-162	KASEYLQLVF
150	MAGE-2 218-225	EEKIWEEL
151	MAGE-2 216-225	APEEKIWEEL
152	MAGE-2 216-223	APEEKIWE
153	MAGE-2 220-228	KIWEELSML
154	MAGE-2 219-228	EKIWEELSML
155	MAGE-2 271-278	FLWGPRAL
156	MAGE-2 271-279	FLWGPRALI
157	MAGE-2 278-286	LIETSYVKV
158	MAGE-2 277-286	ALIETSYVKV
159	MAGE-2 276-284	RALIETSYV

160	MAGE-2 279-287	IETSYVKVL
161	MAGE-2 278-287	LIETSYVKVL
162	MAGE-3 271-278	FLWGPRAL
163	MAGE-3 270-278	EFLWGPRAL
164	MAGE-3 271-279	FLWGPRALV
165	MAGE-3 276-284	RALVETSYV
166	MAGE-3 272-280	LWGPRALVE
167	MAGE-3 271-280	FLWGPRALVE
168	MAGE-3 27 2-281	LWGPRALVET
169	NY-ESO-1 82-90	GPESRLLEF
170	NY-ESO-1 83-91	PESRLLEFY
171	NY-ESO-1 82-91	GPESRLLEFY
172	NY-ESO-1 84-92	ESRLLEFYL
173	NY-ESO-1 86-94	RLLEFYLAM
174	NY-ESO-1 88-96	LEFYLAMPF
175	NY-ESO-1 87-96	LLEFYLAMPF
176	NY-ESO-1 93-102	AMPFATPMEA
177	NY-ESO-1 94-102	MPFATPMEA
178	NY-ESO-1 115-123	PLPVPGVLL
179	NY-ESO-1 114-123	PPLPVPGVLL
180	NY-ESO-1 116-123	LPVPGVLL
181	NY-ESO-1 103-112	ELARRSLAQD
182	NY-ESO-1 118-126	VPGVLLKEF
183	NY-ESO-1 117-126	PVPGVLLKEF
184	NY-ESO-1 116-123	LPVPGVLL
185	NY-ESO-1 127-135	TVSGNILTI
186	NY-ESO-1 126-135	FTVSGNILTI
187	NY-ESO-1 120-128	GVLLKEFTV
188	NY-ESO-1 121-130	VLLKEFTVSG
189	NY-ESO-1 122-130	LLKEFTVSG
190	NY-ESO-1 118-126	VPGVLLKEF
191	NY-ESO-1 117-126	PVPGVLLKEF
192	NY-ESO-1 139-147	AADHRQLQL
193	NY-ESO-1 148-156	SISSCLQQL
194	NY-ESO-1 147-156	LSISSCLQQL
195	NY-ESO-1 138-147	TAADHRQLQL
196	NY-ESO-1 161-169	WITQCFLPV
197	NY-ESO-1 157-165	SLLMWITQC
198	NY-ESO-1 150-158	SSCLQQQLSL
199	NY-ESO-1 154-162	QQLSLLMWI
200	NY-ESO-1 151-159	SCLQQQLSLL
201	NY-ESO-1 150-159	SSCLQQQLSLL

202	NY-ESO-1 163-171	TQCFLPVFL
203	NY-ESO-1 162-171	ITQCFLPVFL
204	PRAME 219-227	PMQDIKMIL
205	PRAME 218-227	MPMQDIKMIL
206	PRAME 428-436	QHLIGLSNL
207	PRAME 427-436	LQHLIGLSNL
208	PRAME 429-436	HLIGLSNL
209	PRAME 431-439	IIGLSNLTHV
210	PRAME 430-439	LIGLSNLTHV
211	PSA 53-61	VLVHPQWVL
212	PSA 52-61	GVLVHPQWVL
213	PSA 52-60	GVLVHPQWVW
214	PSA 59-67	WVLTAAHCI
215	PSA 54-63	LVHPQWVLTA
216	PSA 53-62	VLVHPQWVLT
217	PSA 54-62	LVHPQWVLT
218	PSA 66-73	CIRNKS VI
219	PSA 65-73	HCIRNKS VI
220	PSA 56-64	HPQWVL TAA
221	PSA 63-72	AAHCIRNKS V
222	PSCA 116-123	LLWGP GQL
223	PSCA 115-123	LLLWGP GQL
224	PSCA 114-123	GLLWGP GQL
225	PSCA 99-107	ALQPAAAIL
226	PSCA 98-107	HALQPAAAIL
227	Tyr 128-137	APEKDKFFAY
228	Tyr 129-137	PEKDKFFAY
229	Tyr 130-138	EKDKFFAYL
230	Tyr 131-138	KDKFFAYL
231	Tyr 205-213	PAFLPWHRL
232	Tyr 204-213	APAFLPWHRL
233	Tyr 214-223	FLLRWEQEIQ
234	Tyr 212-220	RLFLLRWEQ
235	Tyr 191-200	GSEIWRDIDF
236	Tyr 192-200	SEIWRDIDF
237	Tyr 473-481	RIWSWLLGA
238	Tyr 476-484	SWLLGAAMV
239	Tyr 477-486	WLLGAAMVGA
240	Tyr 478-486	LLGAAMVGA
241	PSMA 4-12	LLHETDSAV
242	PSMA 13-21	ATARRPRWL
243	PSMA 53-61	TPKHNMKAF

244	PSMA 64-73	ELKAENIKKF
245	PSMA 69-77	NIKKFLH <sup>1</sup> NF
246	PSMA 68-77	ENIKKFLH <sup>1</sup> NF
247	PSMA 220-228	AGAKGVILY
248	PSMA 468-477	PLMYSLVHNL
249	PSMA 469-477	LMYSLVHNL
250	PSMA 463-471	RVDCTPLMY
251	PSMA 465-473	DCTPLMYSL
252	PSMA 507-515	SGMPRISKL
253	PSMA 506-515	FSGMPRISKL
254	NY-ESO-1 136-163	RLTAADHRQLQLSISSCLQQQLSLMWIT
255	NY-ESO-1 150-177	SSCLQQQLSLMWITQCFLPVFLAQPPSG

<sup>1</sup>This H was reported as Y in the SWISSPROT database.

Table 1C. SEQ ID NOS.\* including epitopes in Example14.

SEQ ID NO.	IDENTITY	SEQUENCE
256	Mage-1 125-132	KAEMLESV
257	Mage-1 124-132	TKAEMLESV
258	Mage-1 123-132	VTKAEMLESV
259	Mage-1 128-136	MLESVIKNY
260	Mage-1 127-136	EMLESVIKNY
261	Mage-1 125-133	KAEMLESVI
262	Mage-1 146-153	KASESQL
263	Mage-1 145-153	GKASESQL
264	Mage-1 147-155	ASESQLVF
265	Mage-1 153-161	LVFGIDVKE
266	Mage-1 114-121	LLKYRARE
267	Mage-1 106-113	VADLVGFL
268	Mage-1 105-113	KVADLVGFL
269	Mage-1 107-115	ADLVGFLL
270	Mage-1 106-115	VADLVGFLLL
271	Mage-1 114-123	LLKYRAREPV
272	Mage-3 278-286	LVETSYVKV
273	Mage-3 277-286	ALVETSYVKV
274	Mage-3 285-293	KVLHHMVKI
275	Mage-3 283-291	YVKVLHHMV
276	Mage-3 275-283	PRALVETSY
277	Mage-3 274-283	GPRALVETSY
278	Mage-3 278-287	LVETSYVKVL
279	ED-B 4'-5	TIIPEVPQL
280	ED-B 5'-5	DTIIPEVPQL
281	ED-B 1-10	EVPQLTDLSF
282	ED-B 23-30	TPLNSSTI
283	ED-B 18-25	IGLRWTPL
284	ED-B 17-25	SIGLRWTPL
285	ED-B 25-33	LNSSTIIGY
286	ED-B 24-33	PLNSSTIIGY

287	ED-B 23-31	TPLNSSTII
288	ED-B 31-38	IGYRITVV
289	ED-B 30-38	IIGYRITVV
290	ED-B 29-38	TIIGYRITVV
291	ED-B 31-39	IGYRITVVA
292	ED-B 30-39	IIGYRITVVA
293	CEA 184-191	SLPVSPRL
294	CEA 183-191	QSLPVSPRL
295	CEA 186-193	PVSPRLQL
296	CEA 185-193	LPVSPRLQL
297	CEA 184-193	SLPVSPRLQL
298	CEA 185-192	LPVSPRLQ
299	CEA 192-200	QLSNGNRTL
300	CEA 191-200	LQLSNGNRTL
301	CEA 179-187	WVNNQSLPV
302	CEA 186-194	PVSPRLQLS
303	CEA 362-369	SLPVSPRL
304	CEA 361-369	QSLPVSPRL
305	CEA 364-371	PVSPRLQL
306	CEA 363-371	LPVSPRLQL
307	CEA 362-371	SLPVSPRLQL
308	CEA 363-370	LPVSPRLQ
309	CEA 370-378	QLSNDNRTL
310	CEA 369-378	LQLSNDNRTL
311	CEA 357-365	WVNNQSLPV
312	CEA 360-368	NQSLPVSPR
313	CEA 540-547	SLPVSPRL
314	CEA 539-547	QSLPVSPRL
315	CEA 542-549	PVSPRLQL
316	CEA 541-549	LPVSPRLQL
317	CEA 540-549	SLPVSPRLQL
318	CEA 541-548	LPVSPRLQ
319	CEA 548-556	QLSNGNRTL
320	CEA 547-556	LQLSNGNRTL
321	CEA 535-543	WVNGQSLPV
322	CEA 533-541	LWWVNGQSL
323	CEA 532-541	YLWWVNGQSL
324	CEA 538-546	GQSLPVSPR
325	Her-2 30-37	DMKLRPA
326	Her-2 28-37	GTDMKLRPA
327	Her-2 42-49	HLDMLRHL
328	Her-2 41-49	THLDMLRHL
329	Her-2 40-49	ETHLDMLRHL
330	Her-2 36-43	PASPETHL
331	Her-2 35-43	LPASPETHL
332	Her-2 34-43	RLPASPETHL
333	Her-2 38-46	SPETHLDML
334	Her-2 37-46	ASPETHLDML
335	Her-2 42-50	HLDMLRHLY
336	Her-2 41-50	THLDMLRHLY
337	Her-2 719-726	ELRKVKVL

338	Her-2 718-726	TELRKVKVL
339	Her-2 717-726	ETELRKVKVL
340	Her-2 715-723	LKETELRKV
341	Her-2 714-723	ILKETELRKV
342	Her-2 712-720	MRILKETEL
343	Her-2 711-720	QMRILKETEL
344	Her-2 717-725	ETELRKVKV
345	Her-2 716-725	KETELRKVKV
346	Her-2 706-714	MPNQAQMRI
347	Her-2 705-714	AMPNQAQMRI
348	Her-2 706-715	MPNQAQMRL
349	HER-2 966-973	RPRFRELV
350	HER-2 965-973	CRPRFRELV
351	HER-2 968-976	RFRELVSEF
352	HER-2 967-976	PRFRELVSEF
353	HER-2 964-972	ECRPRFREL
354	NY-ESO-1 67-75	GAASGLNGC
355	NY-ESO-1 52-60	RASGPGGGA
356	NY-ESO-1 64-72	PHGGAASGL
357	NY-ESO-1 63-72	GPHGGAASGL
358	NY-ESO-1 60-69	APRGPHGAA
359	PRAME 112-119	VRPRRWKL
360	PRAME 111-119	EVRPRRWKL
361	PRAME 113-121	RPRRWKLQV
362	PRAME 114-122	PRRWKLQVL
363	PRAME 113-122	PRRWKLQVL
364	PRAME 116-124	RWKLQVLSDL
365	PRAME 115-124	RRWKLQVLSDL
366	PRAME 174-182	PVEVLVDLF
367	PRAME 199-206	VKRKKNVL
368	PRAME 198-206	KVKRKKNVL
369	PRAME 197-206	EKVKRKKNVL
370	PRAME 198-205	KVKRKKNV
371	PRAME 201-208	RKKNVRL
372	PRAME 200-208	KRKKNVRL
373	PRAME 199-208	VKRKKNVRL
374	PRAME 189-196	DELFSYLI
375	PRAME 205-213	VLRLCCKKL
376	PRAME 204-213	NVRLCCKKL
377	PRAME 194-202	YLIEKVKRK
378	PRAME 74-81	QAWPFTCL
379	PRAME 73-81	VQAWPFTCL
380	PRAME 72-81	MVQAWPFTCL
381	PRAME 81-88	LPLGVLMK
382	PRAME 80-88	CLPLGVLMK
383	PRAME 79-88	TCLPLGVLMK
384	PRAME 84-92	GVLMKGQHL
385	PRAME 81-89	LPLGVLMKG
386	PRAME 80-89	CLPLGVLMKG
387	PRAME 76-85	WPFTCLPLGV
388	PRAME 51-59	ELFPPLFMA

389	PRAME 49-57	PRELFPPPLF
390	PRAME 48-57	LPRELFPPPLF
391	PRAME 50-58	RELFPPPLFM
392	PRAME 49-58	PRELFPPPLFM
393	PSA 239-246	RPSLYTKV
394	PSA 238-246	ERPSLYTKV
395	PSA 236-243	LPERPSLY
396	PSA 235-243	ALPERPSLY
397	PSA 241-249	SLYTKVVHY
398	PSA 240-249	PSLYTKVVHY
399	PSA 239-247	RPSLYTKVV
400	PSMA 211-218	GNKVKNAQ
401	PSMA 202-209	IARYGKVF
402	PSMA 217-225	AQLAGAKGV
403	PSMA 207-215	KVFRGNKVK
404	PSMA 211-219	GNKVKNAQL
405	PSMA 269-277	TPGYPANEY
406	PSMA 268-277	LTPGYPANEY
407	PSMA 271-279	GYPANEYAY
408	PSMA 270-279	PGYPANEYAY
409	PSMA 266-274	DPLTPGYPA
410	PSMA 492-500	SLYESWTKK
411	PSMA 491-500	KSLYESWTKK
412	PSMA 486-494	EGFEGKSLY
413	PSMA 485-494	DEGFEGKSLY
414	PSMA 498-506	TKKSPSPEF
415	PSMA 497-506	WTKKSPSPEF
416	PSMA 492-501	SLYESWTKKS
417	PSMA 725-732	WGEVKRQI
418	PSMA 724-732	AWGEVKRQI
419	PSMA 723-732	KAWGEVKRQI
420	PSMA 723-730	KAWGEVKR
421	PSMA 722-730	SKAWGEVKR
422	PSMA 731-739	QIYVAAFTV
423	PSMA 733-741	YVAAFTVQA
424	PSMA 725-733	WGEVKRQIY
425	PSMA 727-735	EVKRQIYVA
426	PSMA 738-746	TVQAAAETL
427	PSMA 737-746	FTVQAAAETL
428	PSMA 729-737	KRQIYVAAF
429	PSMA 721-729	PSKAWGEVK
430	PSMA 723-731	KAWGEVKRQ
431	PSMA 100-108	WKEFGLDSV
432	PSMA 99-108	QWKEFGLDSV
433	PSMA 102-111	FFGLDSVELA
434	SCP-1 126-134	ELRQKESKL
435	SCP-1 125-134	AELRQKESKL
436	SCP-1 133-141	KLQENRKII
437	SCP-1 298-305	QLEEKTKL
438	SCP-1 297-305	NQLEEKTKL
439	SCP-1 288-296	LLEESRDKV

440	SCP-1 287-296	FLLEESRDKV
441	SCP-1 291-299	ESRDKVNQL
442	SCP-1 290-299	EESRDKVNL
443	SCP-1 475-483	EKEVHDLEY
444	SCP-1 474-483	REKEVHDLEY
445	SCP-1 480-488	DLEYSYCHY
446	SCP-1 477-485	EVHDLEYSY
447	SCP-1 477-486	EVHDLEYSYC
448	SCP-1 502-509	KLSSKREL
449	SCP-1 508-515	ELKNTEYF
450	SCP-1 507-515	RELKNTEYF
451	SCP-1 496-503	KRGQRPKL
452	SCP-1 494-503	LPKRGQRPKL
453	SCP-1 509-517	LNKNTFYFTL
454	SCP-1 508-517	ELKNTEYFTL
455	SCP-1 506-514	KRELKNTEY
456	SCP-1 502-510	KLSSKRELK
457	SCP-1 498-506	GQRPKLSSK
458	SCP-1 497-506	RGQRPKLSSK
459	SCP-1 500-508	RPKLSSKRE
460	SCP-1 573-580	LEYVREEL
461	SCP-1 572-580	ELEYVREEL
462	SCP-1 571-580	NELEYVREEL
463	SCP-1 579-587	ELKQKREDEV
464	SCP-1 575-583	YVREELKQK
465	SCP-1 632-640	QLNVYEIKV
466	SCP-1 630-638	SKQLNVYEI
467	SCP-1 628-636	AESKQLNVY
468	SCP-1 627-636	TAESKQLNVY
469	SCP-1 638-645	IKVNKLEL
470	SCP-1 637-645	EIKVNKLEL
471	SCP-1 636-645	YEIKVNKLEL
472	SCP-1 642-650	KLELELESA
473	SCP-1 635-643	VYEIKVNKL
474	SCP-1 634-643	NVYEIKVNKL
475	SCP-1 646-654	ELESAKQKF
476	SCP-1 642-650	KLELELESA
477	SCP-1 646-654	ELESAKQKF
478	SCP-1 771-778	KEKLKREA
479	SCP-1 777-785	EAKENTATL
480	SCP-1 776-785	REAKENTATL
481	SCP-1 773-782	KLKREAKENT
482	SCP-1 112-119	EAEKIKKW
483	SCP-1 101-109	GLSRVYSKL
484	SCP-1 100-109	EGLSRVYSKL
485	SCP-1 108-116	KLYKEAEKI
486	SCP-1 98-106	NSEGLSRVY
487	SCP-1 97-106	ENSEGLSRVY
488	SCP-1 102-110	LSRVYSKLY
489	SCP-1 101-110	GLSRVYSKLY
490	SCP-1 96-105	LENSEGLSRV

491	SCP-1 108-117	KLYKEAEKIK
492	SCP-1 949-956	REDRWAVI
493	SCP-1 948-956	MREDRWAVI
494	SCP-1 947-956	KMREDRWAVI
495	SCP-1 947-955	KMREDRWAV
496	SCP-1 934-942	TTPGSTLKF
497	SCP-1 933-942	LITTPGSTLKF
498	SCP-1 937-945	GSTLKGAI
499	SCP-1 945-953	IRKMREDRW
500	SCP-1 236-243	RLEMHFKL
501	SCP-1 235-243	SRLEMHFKL
502	SCP-1 242-250	KLKEDYEKI
503	SCP-1 249-257	KIQHLEQEY
504	SCP-1 248-257	EKIQHLEQEY
505	SCP-1 233-242	ENSRLEMHF
506	SCP-1 236-245	RLEMHFKLKE
507	SCP-1 324-331	LEDIKVSL
508	SCP-1 323-331	ELEDIKVSL
509	SCP-1 322-331	KELEDIKVSL
510	SCP-1 320-327	LTKELEDI
511	SCP-1 319-327	HLTKELEDI
512	SCP-1 330-338	SLQRSVSTQ
513	SCP-1 321-329	TKELEDIKV
514	SCP-1 320-329	LTKELEDIKV
515	SCP-1 326-335	DIKVSLQRSV
516	SCP-1 281-288	KMKDLTFL
517	SCP-1 280-288	NKMKDLTFL
518	SCP-1 279-288	ENKMKDLTFL
519	SCP-1 288-296	LLEESRDKV
520	SCP-1 287-296	FLLEESRDKV
521	SCP-1 291-299	ESRDKVNQL
522	SCP-1 290-299	EESRDKVNQL
523	SCP-1 277-285	EKENKMKDL
524	SCP-1 276-285	TEKENKMKDL
525	SCP-1 279-287	ENKMKDLTF
526	SCP-1 218-225	IEKMITAF
527	SCP-1 217-225	NIEKMITAF
528	SCP-1 216-225	SNIKMITAF
529	SCP-1 223-230	TAFFELRV
530	SCP-1 222-230	ITAFFELRV
531	SCP-1 221-230	MITAFFELRV
532	SCP-1 220-228	KMITAFFEL
533	SCP-1 219-228	EKMITAFFEL
534	SCP-1 227-235	ELRVQAENS
535	SCP-1 213-222	DLNSNIEKMI
536	SCP-1 837-844	WTSAKNTL
537	SCP-1 846-854	TPLPKAYTV
538	SCP-1 845-854	STPLPKAYTV
539	SCP-1 844-852	LSTPLPKAY
540	SCP-1 843-852	TLSTPLPKAY
541	SCP-1 842-850	NTLSTPLPK

542	SCP-1 841-850	KNTLSTPLPK
543	SCP-1 828-835	ISKDKRDY
544	SCP-1 826-835	HGISKDKRDY
545	SCP-1 832-840	KRDYLWTS
546	SCP-1 829-838	SKDKRDYLWT
547	SCP-1 279-286	ENKMKDLT
548	SCP-1 260-268	EINDKEKQV
549	SCP-1 274-282	QITEKENKM
550	SCP-1 269-277	SLLIQITE
551	SCP-1 453-460	FEKIAEEL
552	SCP-1 452-460	QFEKIAEEL
553	SCP-1 451-460	KQFEKIAEEL
554	SCP-1 449-456	DNKQFEKI
555	SCP-1 448-456	YDNKQFEKI
556	SCP-1 447-456	LYDNKQFEKI
557	SCP-1 440-447	LGEKETLL
558	SCP-1 439-447	VLGEKETLL
559	SCP-1 438-447	KVLGEKETLL
560	SCP-1 390-398	LLRTEQQRL
561	SCP-1 389-398	ELLRTEQQRL
562	SCP-1 393-401	TEQQRLENY
563	SCP-1 392-401	RTEQQRLENY
564	SCP-1 402-410	EDQLIILTM
565	SCP-1 397-406	RLENYEDQLI
566	SCP-1 368-375	KARAHSF
567	SCP-1 376-384	VVTEFETTV
568	SCP-1 375-384	FVVTEFETTV
569	SCP-1 377-385	VTEFETTV
570	SCP-1 376-385	VVTEFETTV
571	SCP-1 344-352	DLQIATNTI
572	SCP-1 347-355	IATNTICQL
573	SCP-1 346-355	QIATNTICQL
574	SSX4 57-65	VMTKLGFKY
575	SSX4 53-61	LNYEVMTKL
576	SSX4 52-61	KLNYEVMTKL
577	SSX4 66-74	TLPPFMRSK
578	SSX4 110-118	KIMPKKPAE
579	SSX4 103-112	SLQRIFPKIM
580	Tyr 463-471	YIKSYLEQA
581	Tyr 459-467	SFQDYIKSY
582	Tyr 458-467	DSFQDYIKSY
583	Tyr 507-514	LPEEKQPL
584	Tyr 506-514	QLPEEKQPL
585	Tyr 505-514	KQLPEEKQPL
586	Tyr 507-515	LPEEKQPLL
587	Tyr 506-515	QLPEEKQPLL
588	Tyr 497-505	SLLCRHKRK
589	ED-B domain of Fibronectin	EVQLTDLSFVDITDSSIGLRWTPLNSSTIIGYRI TVVAAAGEGIPIFEDFVDSSVGYYTVTGLEPGID YDISVITLINGGESAPTTLTQQT
590	ED-B domain of	CTFDNLSPGLEYNVSVYTVKDDKESVPISDTIIP

	Fibronectin with flanking sequence from Fibronectin	EV PQLTDLSFVDITDSSIGLRWTPLNSSTIIGYRI TVVAAGEGIPIFEDFVDSSVGYYTVTGLEPGID YDISVITLINGGESAPTTLTQQT AVPPPTDLRFTNIGPDTMRVTW
591	ED-B domain of Fibronectin cds	Accession number: X07717
592	CEA protein	Accession number: P06731
593	CEA cDNA	Accession number: NM 004363
594	Her2/Neu protein	Accession number: P04626
595	Her2/Neu cDNA	Accession number: M11730
596	SCP-1 protein	Accession number: Q15431
597	SCP-1 cDNA	Accession number: X95654
598	SSX-4 protein	Accession number: O60224
599	SSX-4 cDNA	Accession number: NM 005636

\*Any of SEQ ID NOS. 1, 8, 9, 11-23, 26-29, 32-44, 47-54, 56-63, 66-68 88-253, and 256-588 can be useful as epitopes in any of the various embodiments of the invention. Any of SEQ ID NOS. 10, 30, 31, 45, 46, 55, 64, 65, 69, 254, and 255 can be useful as sequences containing epitopes or epitope clusters, as described in various embodiments of the invention.

5

\*\*All accession numbers used here and throughout can be accessed through the NCBI databases, for example, through the Entrez seek and retrieval system on the world wide web.

10 Note that the following discussion sets forth the inventors' understanding of the operation of the invention. However, it is not intended that this discussion limit the patent to any particular theory of operation not set forth in the claims.

15 In pursuing the development of epitope vaccines others have generated lists of predicted epitopes based on MHC binding motifs. Such peptides can be immunogenic, but may not correspond to any naturally produced antigenic fragment so that whole antigen will not elicit a similar response or sensitize a target cell to cytolysis by CTL. Therefore such lists do not differentiate between those sequences that can be useful as vaccines and those that cannot. Efforts to determine which of these predicted epitopes are in fact naturally produced have often relied on screening their reactivity with tumor infiltrating lymphocytes (TIL). However, TIL are strongly biased to recognize immune epitopes whereas tumors (and chronically infected cells) will generally 20 present housekeeping epitopes. Thus, unless the epitope is produced by both the housekeeping and immuno- proteasomes, the target cell will generally not be recognized by CTL induced with TIL-identified epitopes. The epitopes of the present invention, in contrast, are generated by the action a specified proteasome, indicating that they can be naturally produced, and enabling their appropriate use. The importance of the distinction between housekeeping and immune epitopes to vaccine 25 design is more fully set forth in PCT publication WO 01/82963A2.

The epitopes of the invention include or encode polypeptide fragments of TAAs that are precursors or products of proteasomal cleavage by a housekeeping or immune proteasome, and that have known or predicted affinity for at least one allele of MHC I. In some embodiments, the epitopes include or encode a polypeptide of about 6 to 25 amino acids in length, preferably about 7

to 20 amino acids in length, more preferably about 8 to 15 amino acids in length, and still more preferably 9 or 10 amino acids in length. However, it is understood that the polypeptides can be larger as long as they do not contain sequences that cause the polypeptides to be directed away from the proteasome or to be destroyed by the proteasome. For immune epitopes, if the larger peptides do not contain such sequences, they can be processed in the pAPC by the immune proteasome. Housekeeping epitopes may also be embedded in longer sequences provided that the sequence is adapted to facilitate liberation of the epitope's C-terminus by action of the immunoproteasome. The sequences of these epitopes can be subjected to computer analysis in order to calculate physical, biochemical, immunologic, or molecular genetic properties such as mass, isoelectric point, predicted mobility in electrophoresis, predicted binding to other MHC molecules, melting temperature of nucleic acid probes, reverse translations, similarity or homology to other sequences, and the like.

In constructing the polynucleotides encoding the polypeptide epitopes of the invention, the gene sequence of the associated TAA can be used, or the polynucleotide can be assembled from any of the corresponding codons. For a 10 amino acid epitope this can constitute on the order of 15  $10^6$  different sequences, depending on the particular amino acid composition. While large, this is a distinct and readily definable set representing a minuscule fraction of the  $>10^{18}$  possible polynucleotides of this length, and thus in some embodiments, equivalents of a particular sequence disclosed herein encompass such distinct and readily definable variations on the listed sequence. 20 In choosing a particular one of these sequences to use in a vaccine, considerations such as codon usage, self-complementarity, restriction sites, chemical stability, etc. can be used as will be apparent to one skilled in the art.

The invention contemplates producing peptide epitopes. Specifically these epitopes are derived from the sequence of a TAA, and have known or predicted affinity for at least one allele of 25 MHC I. Such epitopes are typically identical to those produced on target cells or pAPCs.

#### Compositions Containing Active Epitopes

Embodiments of the present invention provide polypeptide compositions, including 30 vaccines, therapeutics, diagnostics, pharmacological and pharmaceutical compositions. The various compositions include newly identified epitopes of TAAs, as well as variants of these epitopes. Other embodiments of the invention provide polynucleotides encoding the polypeptide epitopes of the invention. The invention further provides vectors for expression of the polypeptide epitopes for purification. In addition, the invention provides vectors for the expression of the polypeptide epitopes in an APC for use as an anti-tumor vaccine. Any of the epitopes or antigens, or nucleic acids encoding the same, from Table 1A can be used. Other embodiments relate to 35 methods of making and using the various compositions.

A general architecture for a class I MHC-binding epitope can be described, and has been reviewed more extensively in Madden, D.R. *Annu. Rev. Immunol.* 13:587-622, 1995. Much of the binding energy arises from main chain contacts between conserved residues in the MHC molecule and the N- and C-termini of the peptide. Additional main chain contacts are made but vary among 5 MHC alleles. Sequence specificity is conferred by side chain contacts of so-called anchor residues with pockets that, again, vary among MHC alleles. Anchor residues can be divided into primary and secondary. Primary anchor positions exhibit strong preferences for relatively well-defined sets of amino acid residues. Secondary positions show weaker and/or less well-defined preferences that can often be better described in terms of less favored, rather than more favored, residues. 10 Additionally, residues in some secondary anchor positions are not always positioned to contact the pocket on the MHC molecule at all. Thus, a subset of peptides exists that bind to a particular MHC molecule and have a side chain-pocket contact at the position in question and another subset exists that show binding to the same MHC molecule that does not depend on the conformation the peptide assumes in the peptide-binding groove of the MHC molecule. The C-terminal residue (P<sub>n</sub>) 15 is preferably a primary anchor residue. For many of the better studied HLA molecules (e.g. A2, A68, B27, B7, B35, and B53) the second position (P2) is also an anchor residue. However, central anchor residues have also been observed including P3 and P5 in HLA-B8, as well as P5 and P<sub>n-3</sub> 20 in the murine MHC molecules H-2D<sup>b</sup> and H-2K<sup>b</sup>, respectively. Since more stable binding will generally improve immunogenicity, anchor residues are preferably conserved or optimized in the design of variants, regardless of their position.

Because the anchor residues are generally located near the ends of the epitope, the peptide can buckle upward out of the peptide-binding groove allowing some variation in length. Epitopes ranging from 8-11 amino acids have been found for HLA-A68, and up to 13 amino acids for HLA-A2. In addition to length variation between the anchor positions, single residue truncations and 25 extensions have been reported and the N- and C-termini, respectively. Of the non-anchor residues, some point up out of the groove, making no contact with the MHC molecule but being available to contact the TCR, very often P1, P4, and P<sub>n-1</sub> for HLA-A2. Others of the non-anchor residues can become interposed between the upper edges of the peptide-binding groove and the TCR, contacting both. The exact positioning of these side chain residues, and thus their effects on binding, MHC 30 fine conformation, and ultimately immunogenicity, are highly sequence dependent. For an epitope to be highly immunogenic it must not only promote stable enough TCR binding for activation to occur, but the TCR must also have a high enough off-rate that multiple TCR molecules can interact sequentially with the same peptide-MHC complex (Kalergis, A.M. et al., *Nature Immunol.* 2:229-234, 2001). Thus without further information about the ternary complex, both conservative and 35 non-conservative substitutions at these positions merit consideration when designing variants.

The polypeptide epitope variants can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations. Variants can be derived from substitution, deletion or insertion of one or more amino acids as compared with the native sequence. Amino acid substitutions can be the result of replacing one amino acid with another 5 amino acid having similar structural and/or chemical properties, such as the replacement of a threonine with a serine. Such replacements are referred to as conservative amino acid replacements, and all appropriate conservative amino acid replacements are considered to be embodiments of one invention. Insertions or deletions can optionally be in the range of about 1 to 10 4, preferably 1 to 2, amino acids. It is generally preferable to maintain the "anchor positions" of the peptide which are responsible for binding to the MHC molecule in question. Indeed, 15 immunogenicity of peptides can be improved in many cases by substituting more preferred residues at the anchor positions (Franco, et al., *Nature Immunology*, 1(2):145-150, 2000). Immunogenicity of a peptide can also often be improved by substituting bulkier amino acids for small amino acids found in non-anchor positions while maintaining sufficient cross-reactivity with the original epitope to constitute a useful vaccine. The variation allowed can be determined by routine 20 insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the polypeptide epitope. Because the polypeptide epitope is often 9 amino acids, the substitutions preferably are made to the shortest active epitope, for example, an epitope of 9 amino acids.

25 Variants can also be made by adding any sequence onto the N-terminus of the polypeptide epitope variant. Such N-terminal additions can be from 1 amino acid up to at least 25 amino acids. Because peptide epitopes are often trimmed by N-terminal exopeptidases active in the pAPC, it is understood that variations in the added sequence can have no effect on the activity of the epitope. In preferred embodiments, the amino acid residues between the last upstream proteasomal cleavage site and the N-terminus of the MHC epitope do not include a proline residue. Serwold, T. at al., 30 *Nature Immunol.* 2:644-651, 2001. Accordingly, effective epitopes can be generated from precursors larger than the preferred 9-mer class I motif.

35 Peptides are useful to the extent that they correspond to epitopes actually displayed by MHC I on the surface of a target cell or a pACP. A single peptide can have varying affinities for different MHC molecules, binding some well, others adequately, and still others not appreciably (Table 2). MHC alleles have traditionally been grouped according to serologic reactivity which does not reflect the structure of the peptide-binding groove, which can differ among different alleles of the same type. Similarly, binding properties can be shared across types; groups based on shared binding properties have been termed supertypes. There are numerous alleles of MHC I in the human population; epitopes specific to certain alleles can be selected based on the genotype of the patient.

**Table 2.**  
**Predicted Binding of Tyrosinase<sub>207-216</sub> (SEQ ID NO. 1) to Various MHC types**

MHC I type	*Half time of dissociation (min)
A1	0.05
A*0201	1311.
A*0205	50.4
A3	2.7
A*1101 (part of the A3 supertype)	0.012
A24	6.0
B7	4.0
B8	8.0
B14 (part of the B27 supertype)	60.0
B*2702	0.9
B*2705	30.0
B*3501 (part of the B7 supertype)	2.0
B*4403	0.1
B*5101 (part of the B7 supertype)	26.0
B*5102	55.0
B*5801	0.20
B60	0.40
B62	2.0

\*HLA Peptide Binding Predictions (internet [http://bimac.dcr.nih.gov/molbio/hla\\_bin](http://bimac.dcr.nih.gov/molbio/hla_bin))

5

In further embodiments of the invention, the epitope, as peptide or encoding polynucleotide, can be administered as a vaccine or immunogenic composition, alone or in combination with various adjuvants, carriers, or excipients. It should be noted that although the term vaccine may be used herein, the discussion can be applied and used with any of the other compositions mentioned herein. Particularly advantageous adjuvants include various cytokines and oligonucleotides containing immunostimulatory sequences (as set forth in greater detail in the co-pending applications referenced herein). Additionally the polynucleotide encoded epitope can be contained in a virus (e.g. *vaccinia* or adenovirus) or in a microbial host cell (e.g. *Salmonella* or *Listeria monocytogenes*) which is then used as a vector for the polynucleotide (Dietrich, G. et al. Nat. Biotech. 16:181-185, 1998). Alternatively a pAPC can be transformed, *ex vivo*, to express the epitope, or pulsed with peptide epitope, to be itself administered as a vaccine. To increase efficiency of these processes, the encoded epitope can be carried by a viral or bacterial vector, or complexed with a ligand of a receptor found on pAPC. Similarly the peptide epitope can be complexed with or conjugated to a pAPC ligand. A vaccine can be composed of more than a single epitope.

10

15

20

Particularly advantageous strategies for incorporating epitopes, and combining them with epitope clusters, into a vaccine are disclosed in U.S. Patent Application No. 09/560,465 entitled "EPIPOPE SYNCHRONIZATION IN ANTIGEN PRESENTING CELLS," filed on April 28,

2000. Epitope clusters for use in connection with this invention are disclosed in U.S. Patent Application No. 09/561,571 entitled "EPITOPE CLUSTERS," filed on April 28, 2000.

Preferred embodiments of the present invention are directed to vaccines and methods for causing a pAPC or population of pAPCs to present housekeeping epitopes that correspond to the epitopes displayed on a particular target cell. Any of the epitopes or antigens in Table 1A, can be used for example. In one embodiment, the housekeeping epitope is a TuAA epitope processed by the housekeeping proteasome of a particular tumor type. In another embodiment, the housekeeping epitope is a virus-associated epitope processed by the housekeeping proteasome of a cell infected with a virus. This facilitates a specific T cell response to the target cells. Concurrent expression by the pAPCs of multiple epitopes, corresponding to different induction states (pre- and post-attack), can drive a CTL response effective against target cells as they display either housekeeping epitopes or immune epitopes.

By having both housekeeping and immune epitopes present on the pAPC, this embodiment can optimize the cytotoxic T cell response to a target cell. With dual epitope expression, the pAPCs can continue to sustain a CTL response to the immune-type epitope when the tumor cell switches from the housekeeping proteasome to the immune proteasome with induction by IFN, which, for example, may be produced by tumor-infiltrating CTLs.

In a preferred embodiment, immunization of a patient is with a vaccine that includes a housekeeping epitope. Many preferred TAAs are associated exclusively with a target cell, particularly in the case of infected cells. In another embodiment, many preferred TAAs are the result of deregulated gene expression in transformed cells, but are found also in tissues of the testis, ovaries and fetus. In another embodiment, useful TAAs are expressed at higher levels in the target cell than in other cells. In still other embodiments, TAAs are not differentially expressed in the target cell compare to other cells, but are still useful since they are involved in a particular function of the cell and differentiate the target cell from most other peripheral cells; in such embodiments, healthy cells also displaying the TAA may be collaterally attacked by the induced T cell response, but such collateral damage is considered to be far preferable to the condition caused by the target cell.

A preferred embodiment of the present invention includes a method of administering a vaccine including a housekeeping epitope to induce a therapeutic immune response. The vaccine is administered to a patient in a manner consistent with the standard vaccine delivery protocols that are well known in the art. Methods of administering epitopes of TAAs include, without limitation, transdermal, intranodal, perinodal, oral, intravenous, intradermal, intramuscular, intraperitoneal, and mucosal administration. A particularly useful method of vaccine delivery to elicit a CTL response is disclosed in PCT Publication No. WO 99/01283, entitled "A METHOD OF INDUCING A CTL RESPONSE," filed on July 10, 1998.

Because the epitope synchronization system has utility in inducing a cell mediated immune response, a vaccine to induce a specific T cell response to a target cell is likewise included in a preferred embodiment of the present invention. The vaccine contains a housekeeping epitope in a concentration effective to cause a pAPC or populations of pAPCs to display housekeeping epitopes. Advantageously, the vaccine can include a plurality of housekeeping epitopes or one or more housekeeping epitopes optionally in combination with one or more immune epitopes. Formulations of the vaccine contain peptides and/or nucleic acids in a concentration sufficient to cause pAPCs to present the epitopes. The formulations preferably contain epitopes in a total concentration of about 1 $\mu$ g-1mg/100 $\mu$ l of vaccine preparation. Conventional dosages and dosing for peptide vaccines and/or nucleic acid vaccines can be used with the present invention, and such dosing regimens are well understood in the art. In one embodiment, a single dosage for an adult human may advantageously be from about 1 to about 5000  $\mu$ l of such a composition, administered one time or multiple times, e.g., in 2, 3, 4 or more dosages separated by 1 week, 2 weeks, 1 month, or more. insulin pump delivers 1 ul per hour (lowest frequency) ref intranodal method patent.

The compositions and methods of the invention disclosed herein further contemplate incorporating adjuvants into the formulations in order to enhance the performance of the vaccines. Specifically, the addition of adjuvants to the formulations is designed to enhance the delivery or uptake of the epitopes by the pAPCs. The adjuvants contemplated by the present invention are known by those of skill in the art and include, for example, GMCSF, GCSF, IL-2, IL-12, BCG, tetanus toxoid, osteopontin, and ETA-1.

In some embodiments of the invention, the vaccines can include a recombinant organism, such as a virus, bacterium or parasite, genetically engineered to express an epitope in a host. For example, *Listeria monocytogenes*, a gram-positive, facultative intracellular bacterium, is a potent vector for targeting TAAAs to the immune system. In a preferred embodiment, this vector can be engineered to express a housekeeping epitope to induce therapeutic responses. The normal route of infection of this organism is through the gut and can be delivered orally. In another embodiment, an adenovirus (Ad) vector encoding a housekeeping epitope for a TAA can be used to induce anti-virus or anti-tumor responses. Bone marrow-derived dendritic cells can be transduced with the virus construct and then injected, or the virus can be delivered directly via subcutaneous injection into an animal to induce potent T-cell responses. Another embodiment employs a recombinant vaccinia virus engineered to encode amino acid sequences corresponding to a housekeeping epitope for a TAA. Vaccinia viruses carrying constructs with the appropriate nucleotide substitutions in the form of a minigene construct can direct the expression of a housekeeping epitope, leading to a therapeutic T cell response against the epitope.

The immunization with DNA requires that APCs take up the DNA and express the encoded proteins or peptides. It is possible to encode a discrete class I peptide on the DNA. By

immunizing with this construct, APCs can be caused to express a housekeeping epitope, which is then displayed on class I MHC on the surface of the cell for stimulating an appropriate CTL response. Constructs generally relying on termination of translation or non-proteasomal proteases for generation of proper termini of housekeeping epitopes have been described in U.S. Patent 5 application No. 09/561,572 entitled EXPRESSION VECTORS ENCODING EPITOPES OF TARGET-ASSOCIATED ANTIGENS, filed on April 28, 2000.

It can be desirable to express housekeeping peptides in the context of a larger protein. Processing can be detected even when a small number of amino acids are present beyond the terminus of an epitope. Small peptide hormones are usually proteolytically processed from longer 10 translation products, often in the size range of approximately 60-120 amino acids. This fact has led some to assume that this is the minimum size that can be efficiently translated. In some embodiments, the housekeeping peptide can be embedded in a translation product of at least about 60 amino acids. In other embodiments the housekeeping peptide can be embedded in a translation product of at least about 50, 30, or 15 amino acids.

15 Due to differential proteasomal processing, the immune proteasome of the pAPC produces peptides that are different from those produced by the housekeeping proteasome in peripheral body cells. Thus, in expressing a housekeeping peptide in the context of a larger protein, it is preferably expressed in the APC in a context other than its full length native sequence, because, as a housekeeping epitope, it is generally only efficiently processed from the native protein by the housekeeping proteasome, which is not active in the APC. In order to encode the housekeeping epitope in a DNA sequence encoding a larger protein, it is useful to find flanking areas on either side of the sequence encoding the epitope that permit appropriate cleavage by the immune proteasome in order to liberate that housekeeping epitope. Such a sequence ensuring epitope 20 synchronization is referred to hereinafter as a SYNCHROTOPE™. Altering flanking amino acid residues at the N-terminus and C-terminus of the desired housekeeping epitope can facilitate appropriate cleavage and generation of the housekeeping epitope in the APC. Sequences embedding housekeeping epitopes can be designed *de novo* and screened to determine which can be successfully processed by immune proteasomes to liberate housekeeping epitopes.

30 Alternatively, another strategy is very effective for identifying sequences allowing production of housekeeping epitopes in APC. A contiguous sequence of amino acids can be generated from head to tail arrangement of one or more housekeeping epitopes. A construct expressing this sequence is used to immunize an animal, and the resulting T cell response is evaluated to determine its specificity to one or more of the epitopes in the array. By definition, 35 these immune responses indicate housekeeping epitopes that are processed in the pAPC effectively. The necessary flanking areas around this epitope are thereby defined. The use of flanking regions of about 4-6 amino acids on either side of the desired peptide can provide the necessary

information to facilitate proteasome processing of the housekeeping epitope by the immune proteasome. Therefore, a SYNCHROTOPE™ of approximately 16-22 amino acids can be inserted into, or fused to, any protein sequence effectively to result in that housekeeping epitope being produced in an APC. In alternate embodiments the whole head-to-tail array of epitopes, or just the 5 epitopes immediately adjacent to the correctly processed housekeeping epitope can be similarly transferred from a test construct to a vaccine vector.

In a preferred embodiment, the housekeeping epitopes can be embedded between known immune epitopes, or segments of such, thereby providing an appropriate context for processing. The abutment of housekeeping and immune epitopes can generate the necessary context to enable 10 the immune proteasome to liberate the housekeeping epitope, or a larger fragment, preferably including a correct C-terminus. It can be useful to screen constructs to verify that the desired epitope is produced. The abutment of housekeeping epitopes can generate a site cleavable by the immune proteasome. Some embodiments of the invention employ known epitopes to flank housekeeping epitopes in test substrates; in others, screening as described below are used whether 15 the flanking regions are arbitrary sequences or mutants of the natural flanking sequence, and whether or not knowledge of proteasomal cleavage preferences are used in designing the substrates.

Cleavage at the mature N-terminus of the epitope, while advantageous, is not required, since a variety of N-terminal trimming activities exist in the cell that can generate the mature N-terminus of the epitope subsequent to proteasomal processing. It is preferred that such N-terminal 20 extension be less than about 25 amino acids in length and it is further preferred that the extension have few or no proline residues. Preferably, in screening, consideration is given not only to cleavage at the ends of the epitope (or at least at its C-terminus), but consideration also can be given to ensure limited cleavage within the epitope.

25 Shotgun approaches can be used in designing test substrates and can increase the efficiency of screening. In one embodiment multiple epitopes can be assembled one after the other, with individual epitopes possibly appearing more than once. The substrate can be screened to determine which epitopes can be produced. In the case where a particular epitope is of concern a substrate can be designed in which it appears in multiple different contexts. When a single epitope appearing in 30 more than one context is liberated from the substrate additional secondary test substrates, in which individual instances of the epitope are removed, disabled, or are unique, can be used to determine which are being liberated and truly constitute SYNCHROTOPE™s.

35 Several readily practicable screens exist. A preferred *in vitro* screen utilizes proteasomal digestion analysis, using purified immune proteasomes, to determine if the desired housekeeping epitope can be liberated from a synthetic peptide embodying the sequence in question. The position of the cleavages obtained can be determined by techniques such as mass spectrometry, HPLC, and

N-terminal pool sequencing; as described in greater detail in U. S. Patent Applications entitled METHOD OF EPITOPE DISCOVERY, EPITOPE SYNCHRONIZATION IN ANTIGEN PRESENTING CELLS, two Provisional U. S. Patent Applications entitled EPITOPE SEQUENCES, which are all cited and incorporated by reference above.

5        Alternatively, *in vivo* screens such as immunization or target sensitization can be employed. For immunization a nucleic acid construct capable of expressing the sequence in question is used. Harvested CTL can be tested for their ability to recognize target cells presenting the housekeeping epitope in question. Such targets cells are most readily obtained by pulsing cells expressing the appropriate MHC molecule with synthetic peptide embodying the mature 10      housekeeping epitope. Alternatively, cells known to express housekeeping proteasome and the antigen from which the housekeeping epitope is derived, either endogenously or through genetic engineering, can be used. To use target sensitization as a screen, CTL, or preferably a CTL clone, 15      that recognizes the housekeeping epitope can be used. In this case it is the target cell that expresses the embedded housekeeping epitope (instead of the pAPC during immunization) and it must express immune proteasome. Generally, the target cell can be transformed with an appropriate nucleic acid construct to confer expression of the embedded housekeeping epitope. Loading with a synthetic peptide embodying the embedded epitope using peptide loaded liposomes or a protein transfer reagent such as BIOPORTER™ (Gene Therapy Systems, San Diego, CA) represents an alternative.

20      Additional guidance on nucleic acid constructs useful as vaccines in accordance with the present invention are disclosed in U.S. Patent Application No. 09/561,572 entitled "EXPRESSION VECTORS ENCODING EPITOPES OF TARGET-ASSOCIATED ANTIGENS," filed on April 28, 2000. Further, expression vectors and methods for their design, which are useful in accordance with the present invention are disclosed in U.S. Patent Application No. 60/336,968 (attorney 25      docket number CTLIMM.022PR) entitled "EXPRESSION VECTORS ENCODING EPITOPES OF TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN," filed on 11/7/2001, which is incorporated by reference in its entirety.

30      A preferred embodiment of the present invention includes a method of administering a vaccine including an epitope (or epitopes) to induce a therapeutic immune response. The vaccine is administered to a patient in a manner consistent with the standard vaccine delivery protocols that are known in the art. Methods of administering epitopes of TAAs including, without limitation, transdermal, intranodal, perinodal, oral, intravenous, intradermal, intramuscular, intraperitoneal, and mucosal administration, including delivery by injection, instillation or inhalation. A particularly useful method of vaccine delivery to elicit a CTL response is disclosed in Australian 35      Patent No. 739189 issued January 17, 2002; U.S. Patent Application No. 09/380,534, filed on

September 1, 1999; and a Continuation-in-Part thereof U.S. Patent Application No. 09/776,232 both entitled "A METHOD OF INDUCING A CTL RESPONSE," filed on February 2, 2001.

Reagents Recognizing Epitopes

5 In another aspect of the invention, proteins with binding specificity for the epitope and the epitope-MHC molecule complex are contemplated, as well as the isolated cells by which they can be expressed. In one set of embodiments these reagents take the form of immunoglobulins: polyclonal sera or monoclonal antibodies (mAb), methods for the generation of which are well known in the art. Generation of mAb with specificity for peptide-MHC molecule complexes is known in the art. See, for example, Aharoni et al. *Nature* 351:147-150, 1991; Andersen et al. *Proc. Natl. Acad. Sci. USA* 93:1820-1824, 1996; Dadaglio et al. *Immunity* 6:727-738, 1997; Duc et al. 10 *Int. Immunol.* 5:427-431, 1993; Eastman et al. *Eur. J. Immunol.* 26:385-393, 1996; Engberg et al. *Immunotechnology* 4:273-278, 1999; Porgdor et al. *Immunity* 6:715-726, 1997; Puri et al. *J. Immunol.* 158:2471-2476, 1997; and Polakova, K., et al. *J. Immunol.* 165 342-348, 2000; all of which are hereby incorporated by reference in their entirety.

15 In other embodiments the compositions can be used to induce and generate, *in vivo* and *in vitro*, T-cells specific for any of the epitopes, including those listed in Table 1A, for example. Thus, embodiments also relate to and include isolated T cells, T cell clones, T cell hybridomas, or a protein containing the T cell receptor (TCR) binding domain derived from the cloned gene, as well 20 as a recombinant cell expressing such a protein. Such TCR derived proteins can be simply the extra-cellular domains of the TCR, or a fusion with portions of another protein to confer a desired property or function. One example of such a fusion is the attachment of TCR binding domains to the constant regions of an antibody molecule so as to create a divalent molecule. The construction and activity of molecules following this general pattern have been reported, for example, Plaksin, D. et al. *J. Immunol.* 158:2218-2227, 1997 and Lebowitz, M.S. et al. *Cell Immunol.* 192:175-184, 25 1999, which are hereby incorporated by reference in their entirety. The more general construction and use of such molecules is also treated in U.S. patent 5,830,755 entitled T CELL RECEPTORS AND THEIR USE IN THERAPEUTIC AND DIAGNOSTIC METHODS, which is hereby incorporated by reference in its entirety.

30 The generation of such T cells can be readily accomplished by standard immunization of laboratory animals, and reactivity to human target cells can be obtained by immunizing with human target cells or by immunizing HLA-transgenic animals with the antigen/epitope. For some therapeutic approaches T cells derived from the same species are desirable. While such a cell can be created by cloning, for example, a murine TCR into a human T cell as contemplated above, *in vitro* immunization of human cells offers a potentially faster option. Techniques for *in vitro* 35 immunization, even using naive donors, are known in the field, for example, Stauss et al., *Proc. Natl. Acad. Sci. USA* 89:7871-7875, 1992; Salgaller et al. *Cancer Res.* 55:4972-4979, 1995; Tsai et

al., *J. Immunol.* 158:1796-1802, 1997; and Chung et al., *J. Immunother.* 22:279-287, 1999; which are hereby incorporated by reference in their entirety.

Any of these molecules can be conjugated to enzymes, radiochemicals, fluorescent tags, and toxins, so as to be used in the diagnosis (imaging or other detection), monitoring, and treatment of the pathogenic condition associated with the epitope. Thus a toxin conjugate can be administered to kill tumor cells, radiolabeling can facilitate imaging of epitope positive tumor, an enzyme conjugate can be used in an ELISA-like assay to diagnose cancer and confirm epitope expression in biopsied tissue. In a further embodiment, such T cells as set forth above, following expansion accomplished through stimulation with the epitope and/or cytokines, can be administered to a patient as an adoptive immunotherapy.

#### Reagents Comprising Epitopes

A further aspect of the invention provides isolated epitope-MHC complexes. In a particularly advantageous embodiment of this aspect of the invention, the complexes can be soluble, multimeric proteins such as those described in U. S. Patent No. 5,635,363 (tetramers) or U. S. Patent No. 6,015,884 (Ig-dimers), both of which are hereby incorporated by reference in their entirety. Such reagents are useful in detecting and monitoring specific T cell responses, and in purifying such T cells.

Isolated MHC molecules complexed with epitopic peptides can also be incorporated into planar lipid bilayers or liposomes. Such compositions can be used to stimulate T cells *in vitro* or, in the case of liposomes, *in vivo*. Co-stimulatory molecules (e.g. B7, CD40, LFA-3) can be incorporated into the same compositions or, especially for *in vitro* work, co-stimulation can be provided by anti-co-receptor antibodies (e.g. anti-CD28, anti-CD154, anti-CD2) or cytokines (e.g. IL-2, IL-12). Such stimulation of T cells can constitute vaccination, drive expansion of T cells *in vitro* for subsequent infusion in an immunootherapy, or constitute a step in an assay of T cell function.

The epitope, or more directly its complex with an MHC molecule, can be an important constituent of functional assays of antigen-specific T cells at either an activation or readout step or both. Of the many assays of T cell function current in the art (detailed procedures can be found in standard immunological references such as *Current Protocols in Immunology* 1999 John Wiley & Sons Inc., N.Y., which is hereby incorporated by reference in its entirety) two broad classes can be defined, those that measure the response of a pool of cells and those that measure the response of individual cells. Whereas the former conveys a global measure of the strength of a response, the latter allows determination of the relative frequency of responding cells. Examples of assays measuring global response are cytotoxicity assays, ELISA, and proliferation assays detecting cytokine secretion. Assays measuring the responses of individual cells (or small clones derived from them) include limiting dilution analysis (LDA), ELISPOT, flow cytometric detection of

unsecreted cytokine (described in U.S. Patent No. 5,445,939, entitled "METHOD FOR ASSESSMENT OF THE MONONUCLEAR LEUKOCYTE IMMUNE SYSTEM" and U.S. Patent Nos 5,656,446; and 5,843,689, both entitled "METHOD FOR THE ASSESSMENT OF THE MONONUCLEAR LEUKOCYTE IMMUNE SYSTEM," reagents for which are sold by Becton, Dickinson & Company under the tradename 'FASTIMMUNE', which patents are hereby incorporated by reference in their entirety) and detection of specific TCR with tetramers or Ig-dimers as stated and referenced above. The comparative virtues of these techniques have been reviewed in Yee, C. et al. *Current Opinion in Immunology*, 13:141-146, 2001, which is hereby incorporated by reference in its entirety. Additionally detection of a specific TCR rearrangement or expression can be accomplished through a variety of established nucleic acid based techniques, particularly in situ and single-cell PCR techniques, as will be apparent to one of skill in the art.

These functional assays are used to assess endogenous levels of immunity, response to an immunologic stimulus (e.g. a vaccine), and to monitor immune status through the course of a disease and treatment. Except when measuring endogenous levels of immunity, any of these assays presume a preliminary step of immunization, whether *in vivo* or *in vitro* depending on the nature of the issue being addressed. Such immunization can be carried out with the various embodiments of the invention described above or with other forms of immunogen (e.g., pAPC-tumor cell fusions) that can provoke similar immunity. With the exception of PCR and tetramer/Ig-dimer type analyses which can detect expression of the cognate TCR, these assays generally benefit from a step of *in vitro* antigenic stimulation which can advantageously use various embodiments of the invention as described above in order to detect the particular functional activity (highly cytolytic responses can sometimes be detected directly). Finally, detection of cytolytic activity requires epitope-displaying target cells, which can be generated using various embodiments of the invention. The particular embodiment chosen for any particular step depends on the question to be addressed, ease of use, cost, and the like, but the advantages of one embodiment over another for any particular set of circumstances will be apparent to one of skill in the art.

#### Tumor Associated Antigens

Epitopes of the present invention are derived from the TuAAs tyrosinase (SEQ ID NO. 2), SSX-2, (SEQ ID NO. 3), PSMA (prostate-specific membrane antigen) (SEQ ID NO. 4), GP100, (SEQ ID NO. 70), MAGE-1, (SEQ ID NO. 71), MAGE-2, (SEQ ID NO. 72), MAGE-3, (SEQ ID NO. 73), NY-ESO-1, (SEQ ID NO. 74), PRAME, (SEQ ID NO. 77), PSA, (SEQ ID NO. 78), and PSCA, (SEQ ID NO. 79). The natural coding sequences for these eleven proteins, or any segments within them, can be determined from their cDNA or complete coding (cds) sequences, SEQ ID NOS. 5-7, and 80-87, respectively.

Tyrosinase is a melanin biosynthetic enzyme that is considered one of the most specific markers of melanocytic differentiation. Tyrosinase is expressed in few cell types, primarily in

melanocytes, and high levels are often found in melanomas. The usefulness of tyrosinase as a TuAA is taught in U.S. Patent 5,747,271 entitled "METHOD FOR IDENTIFYING INDIVIDUALS SUFFERING FROM A CELLULAR ABNORMALITY SOME OF WHOSE ABNORMAL CELLS PRESENT COMPLEXES OF HLA-A2/TYROSINASE DERIVED PEPTIDES, AND METHODS FOR TREATING SAID INDIVIDUALS" which is hereby incorporated by reference in its entirety.

5 GP100, also known as PMel17, also is a melanin biosynthetic protein expressed at high levels in melanomas. GP100 as a TuAA is disclosed in U.S. Patent 5,844,075 entitled "MELANOMA ANTIGENS AND THEIR USE IN DIAGNOSTIC AND THERAPEUTIC METHODS," which is hereby incorporated by reference in its entirety.

10 SSX-2, also known as Hom-Mel-40, is a member of a family of highly conserved cancer-testis antigens (Gure, A.O. et al. *Int. J. Cancer* 72:965-971, 1997, which is hereby incorporated by reference in its entirety). Its identification as a TuAA is taught in U.S. Patent 6,025,191 entitled "ISOLATED NUCLEIC ACID MOLECULES WHICH ENCODE A MELANOMA SPECIFIC 15 ANTIGEN AND USES THEREOF," which is hereby incorporated by reference in its entirety. Cancer-testis antigens are found in a variety of tumors, but are generally absent from normal adult tissues except testis. Expression of different members of the SSX family have been found variously in tumor cell lines. Due to the high degree of sequence identity among SSX family members, similar epitopes from more than one member of the family will be generated and able to 20 bind to an MHC molecule, so that some vaccines directed against one member of this family can cross-react and be effective against other members of this family (see example 3 below).

25 MAGE-1, MAGE-2, and MAGE-3 are members of another family of cancer-testis antigens originally discovered in melanoma (MAGE is a contraction of melanoma-associated antigen) but found in a variety of tumors. The identification of MAGE proteins as TuAAs is taught in U.S. Patent 5,342,774 entitled NUCLEOTIDE SEQUENCE ENCODING THE TUMOR REJECTION ANTIGEN PRECURSOR, MAGE-1, which is hereby incorporated by reference in its entirety, and in numerous subsequent patents. Currently there are 17 entries for (human) MAGE in the SWISS 30 Protein database. There is extensive similarity among these proteins so in many cases, an epitope from one can induce a cross-reactive response to other members of the family. A few of these have not been observed in tumors, most notably MAGE-H1 and MAGE-D1, which are expressed in testes and brain, and bone marrow stromal cells, respectively. The possibility of cross-reactivity on normal tissue is ameliorated by the fact that they are among the least similar to the other MAGE proteins.

35 NY-ESO-1, is a cancer-testis antigen found in a wide variety of tumors, also known as CTAG-1 (Cancer-Testis Antigen-1) and CAG-3 (Cancer Antigen-3). NY-ESO-1 as a TuAA is disclosed in U.S. Patent 5,804,381 entitled ISOLATED NUCLEIC ACID MOLECULE

ENCODING AN ESOPHAGEAL CANCER ASSOCIATED ANTIGEN, THE ANTIGEN ITSELF, AND USES THEREOF which is hereby incorporated by reference in its entirety. A paralogous locus encoding antigens with extensive sequence identity, LAGE-1a/s (SEQ ID NO. 75) and LAGE-1b/L (SEQ ID NO. 76), have been disclosed in publicly available assemblies of the human genome, and have been concluded to arise through alternate splicing. Additionally, CT-2 (or CTAG-2, Cancer-Testis Antigen-2) appears to be either an allele, a mutant, or a sequencing 5 discrepancy of LAGE-1b/L. Due to the extensive sequence identity, many epitopes from NY-ESO-1 can also induce immunity to tumors expressing these other antigens. See figure 1. The proteins are virtually identical through amino acid 70. From 71-134 the longest run of identities between 10 NY-ESO-1 and LAGE is 6 residues, but potentially cross-reactive sequences are present. And from 135-180 NY-ESO and LAGE-1a/s are identical except for a single residue, but LAGE-1b/L is unrelated due to the alternate splice. The CAMEL and LAGE-2 antigens appear to derive from the LAGE-1 mRNA, but from alternate reading frames, thus giving rise to unrelated protein sequences. More recently, GenBank Accession AF277315.5, Homo sapiens chromosome X clone RP5- 15 865E18, RP5-1087L19, complete sequence, reports three independent loci in this region which are labeled as LAGE1 (corresponding to CTAG-2 in the genome assemblies), plus LAGE2-A and LAGE2-B (both corresponding to CTAG-1 in the genome assemblies).

PSMA (prostate-specific membranes antigen), a TuAA described in U.S. Patent 5,538,866 entitled "PROSTATE-SPECIFIC MEMBRANES ANTIGEN" which is hereby incorporated by 20 reference in its entirety, is expressed by normal prostate epithelium and, at a higher level, in prostatic cancer. It has also been found in the neovasculature of non-prostatic tumors. PSMA can thus form the basis for vaccines directed to both prostate cancer and to the neovasculature of other tumors. This later concept is more fully described in a provisional U.S. Patent application No. 60/274,063 entitled ANTI-NEOVASCULAR VACCINES FOR CANCER, filed March 7, 2001, and U.S. Application No. \_\_\_\_/\_\_\_\_\_, attorney docket number CTLIMM.015A, filed on March 7, 25 2002, entitled "ANTI-NEOVASCULAR PREPARATIONS FOR CANCER," both of which are hereby incorporated by reference in their entirety. Alternate splicing of the PSMA mRNA also leads to a protein with an apparent start at Met<sub>58</sub>, thereby deleting the putative membrane anchor region of PSMA as described in U.S. Patent 5,935,818 entitled "ISOLATED NUCLEIC ACID 30 MOLECULE ENCODING ALTERNATIVELY SPLICED PROSTATE-SPECIFIC MEMBRANES ANTIGEN AND USES THEREOF" which is hereby incorporated by reference in its entirety. A protein termed PSMA-like protein, Genbank accession number AF261715, is nearly identical to amino acids 309-750 of PSMA and has a different expression profile. Thus the most preferred epitopes are those with an N-terminus located from amino acid 58 to 308.

35 PRAME, also known as MAPE, DAGE, and OIP4, was originally observed as a melanoma antigen. Subsequently, it has been recognized as a CT antigen, but unlike many CT antigens (e.g.,

MAGE, GAGE, and BAGE) it is expressed in acute myeloid leukemias. PRAME is a member of the MAPE family which consists largely of hypothetical proteins with which it shares limited sequence similarity. The usefulness of PRAME as a TuAA is taught in U.S. Patent 5,830,753 entitled "ISOLATED NUCLEIC ACID MOLECULES CODING FOR TUMOR REJECTION ANTIGEN PRECURSOR DAGE AND USES THEREOF" which is hereby incorporated by reference in its entirety.

PSA, prostate specific antigen, is a peptidase of the kallikrein family and a differentiation antigen of the prostate. Expression in breast tissue has also been reported. Alternate names include gamma-seminoprotein, kallikrein 3, seminogelase, seminin, and P-30 antigen. PSA has a high degree of sequence identity with the various alternate splicing products prostatic/glandular kallikrein-1 and -2, as well as kallikrein 4, which is also expressed in prostate and breast tissue. Other kallikreins generally share less sequence identity and have different expression profiles. Nonetheless, cross-reactivity that might be provoked by any particular epitope, along with the likelihood that that epitope would be liberated by processing in non-target tissues (most generally by the housekeeping proteasome), should be considered in designing a vaccine.

PSCA, prostate stem cell antigen, and also known as SCAH-2, is a differentiation antigen preferentially expressed in prostate epithelial cells, and overexpresssed in prostate cancers. Lower level expression is seen in some normal tissues including neuroendocrine cells of the digestive tract and collecting ducts of the kidney. PSCA is described in U.S. Patent 5,856,136 entitled "HUMAN STEM CELL ANTIGENS" which is hereby incorporated by reference in its entirety.

Synaptonemal complex protein 1 (SCP-1), also known as HOM-TES-14, is a meiosis-associated protein and also a cancer-testis antigen (Tureci, O., et al. *Proc. Natl. Acad. Sci. USA* 95:5211-5216, 1998). As a cancer antigen its expression is not cell-cycle regulated and it is found frequently in gliomas, breast, renal cell, and ovarian carcinomas. It has some similarity to myosins, but with few enough identities that cross-reactive epitopes are not an immediate prospect.

The ED-B domain of fibronectin is also a potential target. Fibronectin is subject to developmentally regulated alternative splicing, with the ED-B domain being encoded by a single exon that is used primarily in oncofetal tissues (Matsuura, H. and S. Hakomori *Proc. Natl. Acad. Sci. USA* 82:6517-6521, 1985; Carnemolla, B. et al. *J. Cell Biol.* 108:1139-1148, 1989; Lordin-Rosa, B. et al. *Cancer Res.* 50:1608-1612, 1990; Nicolo, G. et al. *Cell Differ. Dev.* 32:401-408, 1990; Borsi, L. et al. *Exp. Cell Res.* 199:98-105, 1992; Oyama, F. et al. *Cancer Res.* 53:2005-2011, 1993; Mandel, U. et al. *APMIS* 102:695-702, 1994; Farnoud, M.R. et al. *Int. J. Cancer* 61:27-34, 1995; Pujuguet, P. et al. *Am. J. Pathol.* 148:579-592, 1996; Gabler, U. et al. *Heart* 75:358-362, 1996; Chevalier, X. *Br. J. Rheumatol.* 35:407-415, 1996; Midulla, M. *Cancer Res.* 60:164-169, 2000).

The ED-B domain is also expressed in fibronectin of the neovasculature (Kaczmarek, J. et al. *Int. J. Cancer* 59:11-16, 1994; Castellani, P. et al. *Int. J. Cancer* 59:612-618, 1994; Neri, D. et al. *Nat. Biotech.* 15:1271-1275, 1997; Karelina, T.V. and A.Z. Eisen *Cancer Detect. Prev.* 22:438-444, 1998; Tarli, L. et al. *Blood* 94:192-198, 1999; Castellani, P. et al. *Acta Neurochir. (Wien)* 142:277-282, 2000). As an oncofetal domain, the ED-B domain is commonly found in the fibronectin expressed by neoplastic cells in addition to being expressed by the neovasculature. Thus, CTL-inducing vaccines targeting the ED-B domain can exhibit two mechanisms of action: direct lysis of tumor cells, and disruption of the tumor's blood supply through destruction of the tumor-associated neovasculature. As CTL activity can decay rapidly after withdrawal of vaccine, interference with normal angiogenesis can be minimal. The design and testing of vaccines targeted to neovasculature is described in Provisional U.S. Patent Application No. 60/274,063 entitled "ANTI-NEOVASCULATURE VACCINES FOR CANCER" and in U.S. Patent Application No. \_\_\_\_\_/\_\_\_\_\_, attorney docket number CTLIMM.015A, entitled "ANTI-NEOVASCULATURE PREPARATIONS FOR CANCER, filed on date even with this application (March 7, 2002). A tumor cell line is disclosed in Provisional U.S. Application No. \_\_\_\_\_/\_\_\_\_\_, filed on March 7, 2002, attorney docket number CTLIMM.028PR, entitled "HLA-TRANSGENIC MURINE TUMOR CELL LINE," which is hereby incorporated by reference in its entirety.

Carcinoembryonic antigen (CEA) is a paradigmatic oncofetal protein first described in 1965 (Gold and Freedman, *J. Exp. Med.* 121: 439-462, 1965. Fuller references can be found in the Online Medelian Inheritance in Man; record \*114890). It has officially been renamed carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5). Its expression is most strongly associated with adenocarcinomas of the epithelial lining of the digestive tract and in fetal colon. CEA is a member of the immunoglobulin supergene family and the defining member of the CEA subfamily.

HER2/NEU is an oncogene related to the epidermal growth factor receptor (van de Vijver, et al., *New Eng. J. Med.* 319:1239-1245, 1988), and apparently identical to the c-ERBB2 oncogene (Di Fiore, et al., *Science* 237: 178-182, 1987). The over-expression of ERBB2 has been implicated in the neoplastic transformation of prostate cancer. As HER2 it is amplified and over-expressed in 25-30% of breast cancers among other tumors where expression level is correlated with the aggressiveness of the tumor (Slamon, et al., *New Eng. J. Med.* 344:783-792, 2001). A more detailed description is available in the Online Medelian Inheritance in Man; record \*164870.

All references mentioned herein are hereby incorporated by reference in their entirety. Further, incorporated by reference in its entirety is U.S. Patent Application No. 10/005,905 (attorney docket number CTLIMM.021CP1) entitled "EPITOPE SYNCHRONIZATION IN ANTIGEN PRESENTING CELLS," filed on November 7, 2001 and a continuation thereof, U.S. Application No. \_\_\_\_\_/\_\_\_\_\_, filed on December 7, 2000, attorney docket number

CTLIMM.21CP1C, also entitled "EPITOPE SYNCHRONIZATION IN ANTIGEN PRESENTING CELLS."

Useful epitopes were identified and tested as described in the following examples. However, these examples are intended for illustration purposes only, and should not be construed as limiting the scope of the invention in any way.

**EXAMPLES**

**Sequences of Specific Preferred Epitopes**

**Example 1**

10 **Manufacture of tyrosinase epitopes.**

A. **Synthetic production of epitopes**

Peptides having an amino acid sequence of any of SEQ ID NO: 1, 8, 9, 11-23, 2-29, 32-44, 47-54, 56-63, 66-68 88-253, or 256-588 are synthesized using either Fmoc or tBOC solid phase synthesis methodologies. After synthesis, the peptides are cleaved from their supports with either trifluoroacetic acid or hydrogen fluoride, respectively, in the presence of appropriate protective scavengers. After removing the acid by evaporation, the peptides are extracted with ether to remove the scavengers and the crude, precipitated peptide is then lyophilized. Purity of the crude peptides is determined by HPLC, sequence analysis, amino acid analysis, counterion content analysis and other suitable means. If the crude peptides are pure enough (greater than or equal to about 90% pure), they can be used as is. If purification is required to meet drug substance specifications, the peptides are purified using one or a combination of the following: re-precipitation; reverse-phase, ion exchange, size exclusion or hydrophobic interaction chromatography; or counter-current distribution.

**Drug product formulation**

25 GMP-grade peptides are formulated in a parenterally acceptable aqueous, organic, or aqueous-organic buffer or solvent system in which they remain both physically and chemically stable and biologically potent. Generally, buffers or combinations of buffers or combinations of buffers and organic solvents are appropriate. The pH range is typically between 6 and 9. Organic modifiers or other excipients can be added to help solubilize and stabilize the peptides. These include detergents, lipids, co-solvents, antioxidants, chelators and reducing agents. In the case of a lyophilized product, sucrose or mannitol or other lyophilization aids can be added. Peptide solutions are sterilized by membrane filtration into their final container-closure system and either lyophilized for dissolution in the clinic, or stored until use.

**B. Construction of expression vectors for use as nucleic acid vaccines**

35 The construction of three generic epitope expression vectors is presented below. The particular advantages of these designs are set forth in U.S. Patent Application No. 09/561,572

entitled "EXPRESSION VECTORS ENCODING EPITOPEs OF TARGET-ASSOCIATED ANTIGENS," which has been incorporated by reference in its entirety above.

5 A suitable *E. coli* strain was then transfected with the plasmid and plated out onto a selective medium. Several colonies were grown up in suspension culture and positive clones were identified by restriction mapping. The positive clone was then grown up and aliquotted into storage vials and stored at -70°C.

A mini-prep (QIAprep Spin Mini-prep: Qiagen, Valencia, CA) of the plasmid was then made from a sample of these cells and automated fluorescent dideoxy sequence analysis was used to confirm that the construct had the desired sequence.

10 B.1 Construction of pVAX-EP1-IRES-EP2

Overview:

The starting plasmid for this construct is pVAX1 purchased from Invitrogen (Carlsbad, CA). Epitopes EP1 and EP2 were synthesized by GIBCO BRL (Rockville, MD). The IRES was excised from pIRES purchased from Clontech (Palo Alto, CA).

15 Procedure:

- 1 pIRES was digested with EcoRI and NotI. The digested fragments were separated by agarose gel electrophoresis, and the IRES fragment was purified from the excised band.
- 2 pVAX1 was digested with EcoRI and NotI, and the pVAX1 fragment was gel-purified.
- 3 The purified pVAX1 and IRES fragments were then ligated together.
- 20 4 Competent *E. coli* of strain DH5 $\alpha$  were transformed with the ligation mixture.
- 5 Minipreps were made from 4 of the resultant colonies.
- 6 Restriction enzyme digestion analysis was performed on the miniprep DNA. One recombinant colony having the IRES insert was used for further insertion of EP1 and EP2. This intermediate construct was called pVAX-IRES.
- 25 7 Oligonucleotides encoding EP1 and EP2 were synthesized.
- 8 EP1 was subcloned into pVAX-IRES between AflII and EcoRI sites, to make pVAX-EP1-IRES;
- 9 EP2 was subcloned into pVAX-EP1-IRES between SalI and NotI sites, to make the final construct pVAX-EP1-IRES-EP2.
- 30 10 The sequence of the EP1-IRES-EP2 insert was confirmed by DNA sequencing.

B.2. Construction of pVAX-EP1-IRES-EP2-ISS-NIS

Overview:

35 The starting plasmid for this construct was pVAX-EP1-IRES-EP2 (Example 1). The ISS (immunostimulatory sequence) introduced into this construct is AACGTT, and the NIS (standing for nuclear import sequence) used is the SV40 72bp repeat sequence. ISS-NIS was synthesized by GIBCO BRL. See Figure 2.

**Procedure:**

- 1 pVAX-EP1-IRES-EP2 was digested with NruI; the linearized plasmid was gel-purified.
- 2 ISS-NIS oligonucleotide was synthesized.
- 3 The purified linearized pVAX-EP1-IRES-EP2 and synthesized ISS-NIS were ligated
- 5 together.
- 4 Competent E. coli of strain DH5 $\alpha$  were transformed with the ligation product.
- 5 Minipreps were made from resultant colonies.
- 6 Restriction enzyme digestions of the minipreps were carried out.
- 7 The plasmid with the insert was sequenced.

**10 B3. Construction of pVAX-EP2-UB-EP1****Overview:**

The starting plasmid for this construct was pVAX1 (Invitrogen). EP2 and EP1 were synthesized by GIBCO BRL. Wild type Ubiquitin cDNA encoding the 76 amino acids in the construct was cloned from yeast.

**15 Procedure:**

- 1 RT-PCR was performed using yeast mRNA. Primers were designed to amplify the complete coding sequence of yeast Ubiquitin.
- 2 The RT-PCR products were analyzed using agarose gel electrophoresis. A band with the predicted size was gel-purified.
- 3 The purified DNA band was subcloned into pZERO1 at EcoRV site. The resulting clone
- 20 was named pZERO-UB.
- 4 Several clones of pZERO-UB were sequenced to confirm the Ubiquitin sequence before further manipulations.
- 5 EP1 and EP2 were synthesized.
- 6 EP2, Ubiquitin and EP1 were ligated and the insert cloned into pVAX1 between BamHI and EcoRI, putting it under control of the CMV promoter.
- 7 The sequence of the insert EP2-UB-EP1 was confirmed by DNA sequencing.

**Example 2****Identification of useful epitope variants.**

30 The 10-mer FLPWHRLFLL (SEQ ID NO. 1) is identified as a useful epitope. Based on this sequence, numerous variants are made. Variants exhibiting activity in HLA binding assays (see Example 3, section 6) are identified as useful, and are subsequently incorporated into vaccines.

The HLA-A2 binding of length variants of FLPWHRLFLL have been evaluated. Proteasomal digestion analysis indicates that the C-terminus of the 9-mer FLPWHRLFL (SEQ ID NO. 8) is also produced. Additionally the 9-mer LPWHRLFLL (SEQ ID NO. 9) can result from N-terminal trimming of the 10-mer. Both are predicted to bind to the HLA-A\*0201 molecule,

however of these two 9-mers, FLPWHRLFL displayed more significant binding and is preferred (see Figs. 3A and B).

Sequence variants of FLPWHRLFL are constructed as follow. Consistent with the binding coefficient table (see Table 3) from the NIH/BIMAS MHC binding prediction program (see reference in example 3 below), binding can be improved by changing the L at position 9, an anchor position, to V. Binding can also be altered, though generally to a lesser extent, by changes at non-anchor positions. Referring generally to Table 3, binding can be increased by employing residues with relatively larger coefficients. Changes in sequence can also alter immunogenicity independently of their effect on binding to MHC. Thus binding and/or immunogenicity can be improved as follows:

By substituting F,L,M,W, or Y for P at position 3; these are all bulkier residues that can also improve immunogenicity independent of the effect on binding. The amine and hydroxyl-bearing residues, Q and N; and S and T; respectively, can also provoke a stronger, cross-reactive response.

By substituting D or E for W at position 4 to improve binding; this addition of a negative charge can also make the epitope more immunogenic, while in some cases reducing cross-reactivity with the natural epitope. Alternatively the conservative substitutions of F or Y can provoke a cross-reactive response.

By substituting F for H at position 5 to improve binding. H can be viewed as partially charged, thus in some cases the loss of charge can hinder cross-reactivity. Substitution of the fully charged residues R or K at this position can enhance immunogenicity without disrupting charge-dependent cross-reactivity.

By substituting I, L, M, V, F, W, or Y for R at position 6. The same caveats and alternatives apply here as at position 5.

By substituting W or F for L at position 7 to improve binding. Substitution of V, I, S, T, Q, or N at this position are not generally predicted to reduce binding affinity by this model (the NIH algorithm), yet can be advantageous as discussed above.

Y and W, which are equally preferred as the Fs at positions 1 and 8, can provoke a useful cross-reactivity. Finally, while substitutions in the direction of bulkiness are generally favored to improve immunogenicity, the substitution of smaller residues such as A, S, and C, at positions 3-7 can be useful according to the theory that contrast in size, rather than bulkiness per se, is an important factor in immunogenicity. The reactivity of the thiol group in C can introduce other properties as discussed in Chen, J.-L., et al. *J. Immunol.* 165:948-955, 2000.

**Table 3. 9-mer Coefficient Table for HLA-A\*0201\***

HLA Coefficient table for file "A_0201_standard"									
Amino Acid Type	1 <sup>st</sup>	2 <sup>nd</sup>	3rd	4th	5th	6th	7th	8th	9th
A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.00
C	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	1.00
D	0.075	0.100	0.400	4.100	1.000	1.000	0.490	1.000	0.00
E	0.075	1.400	0.064	4.100	1.000	1.000	0.490	1.000	0.00
F	4.600	0.050	3.700	1.000	3.800	1.900	5.800	5.500	0.01
G	1.000	0.470	1.000	1.000	1.000	1.000	0.130	1.000	0.01
H	0.034	0.050	1.000	1.000	1.000	1.000	1.000	1.000	0.01
I	1.700	9.900	1.000	1.000	1.000	2.300	1.000	0.410	2.10
K	3.500	0.100	0.035	1.000	1.000	1.000	1.000	1.000	0.00
L	1.700	72.000	3.700	1.000	1.000	2.300	1.000	1.000	4.30
M	1.700	52.000	3.700	1.000	1.000	2.300	1.000	1.000	1.00
N	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.01
P	0.022	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.00
Q	1.000	7.300	1.000	1.000	1.000	1.000	1.000	1.000	0.00
R	1.000	0.010	0.076	1.000	1.000	1.000	0.200	1.000	0.00
S	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.01
T	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.50
V	1.700	6.300	1.000	1.000	1.000	2.300	1.000	0.410	14.00
W	4.600	0.010	8.300	1.000	1.000	1.700	7.500	5.500	0.01
Y	4.600	0.010	3.200	1.000	1.000	1.500	1.000	5.500	0.01

\*This table and other comparable data that are publicly available are useful in designing epitope variants and in determining whether a particular variant is substantially similar, or is functionally similar.

5

### Example 3

#### Cluster Analysis (SSX-2<sub>31-69</sub>).

##### 1. Epitope cluster region prediction:

The computer algorithms: SYFPEITHI (internet <http://syfpeithi.bmi-heidelberg.com/Scripts/MHCServer.dll/EpPredict.htm>), based on the book "MHC Ligands and Peptide Motifs" by H.G.Rammensee, J.Bachmann and S.Stevanovic; and HLA Peptide Binding Predictions (NIH) (internet [http://bimas.dcrt.nih.gov/molbio/hla\\_bin](http://bimas.dcrt.nih.gov/molbio/hla_bin)), described in Parker, K. C., et al., *J. Immunol.* 152:163, 1994; were used to analyze the protein sequence of SSX-2 (GI:10337583). Epitope clusters (regions with higher than average density of peptide fragments with high predicted MHC affinity) were defined as described fully in U.S. Patent Application No. 09/561,571 entitled "EPITOPE CLUSTERS," filed on April 28, 2000. Using a epitope density ratio cutoff of 2, five and two clusters were defined using the SYFPEITHI and NIH algorithms, respectively, and peptides score cutoffs of 16 (SYFPEITHI) and 5 (NIH). The highest scoring peptide with the NIH algorithm, SSX-2<sub>41-49</sub>, with an estimated halftime of dissociation of

>1000 min., does not overlap any other predicted epitope but does cluster with SSX-2<sub>57-65</sub> in the NIH analysis.

2. Peptide synthesis and characterization:

5 SSX-2<sub>31-68</sub>, YFSKEEWEKMKASEKIFYVYVMRKYEAATKLGFKATLP (SEQ ID NO. 10) was synthesized by MPS (Multiple Peptide Systems, San Diego, CA 92121) using standard solid phase chemistry. According to the provided 'Certificate of Analysis', the purity of this peptide was 95%.

3. Proteasome digestion:

10 Proteasome was isolated from human red blood cells using the proteasome isolation protocol described in U.S. Patent Application No. 09/561,074 entitled "METHOD OF EPITOPE DISCOVERY," filed on April 28, 2000. SDS-PAGE, western-blotting, and ELISA were used as quality control assays. The final concentration of proteasome was 4 mg/ml, which was determined by non-interfering protein assay (Geno Technologies Inc.). Proteasomes were stored at -70°C in 25 µL aliquots.

15 SSX-2<sub>31-68</sub> was dissolved in Milli-Q water, and a 2 mM stock solution prepared and 20 µL aliquots stored at -20°C.

20 1 tube of proteasome (25 µL) was removed from storage at -70°C and thawed on ice. It was then mixed thoroughly with 12.5 µL of 2mM peptide by repipetting (samples were kept on ice). A 5 µL sample was immediately removed after mixing and transferred to a tube containing 1.25 µL 10%TFA (final concentration of TFA was 2%); the T=0 min sample. The proteasome digestion reaction was then started and carried out at 37°C in a programmable thermal controller. Additional 5 µL samples were taken out at 15, 30, 60, 120, 180 and 240 min respectively, the reaction was stopped by adding the sample to 1.25 µL 10% TFA as before. Samples were kept on ice or frozen until being analyzed by MALDI-MS. All samples were saved and stored at -20°C for HPLC analysis and N-terminal sequencing. Peptide alone (without proteasome) was used as a blank control: 2 µL peptide + 4 µL Tris buffer (20 mM, pH 7.6) + 1.5 µL TFA.

25 4. MALDI-TOF MS measurements:

30 For each time point 0.3 µL of matrix solution (10mg/ml α-cyano-4-hydroxycinnamic acid in AcCN/H<sub>2</sub>O (70:30)) was first applied on a sample slide, and then an equal volume of digested sample was mixed gently with matrix solution on the slide. The slide was allowed to dry at ambient air for 3-5 min. before acquiring the mass spectra. MS was performed on a Lasermat 2000 MALDI-TOF mass spectrometer that was calibrated with peptide/protein standards. To improve the accuracy of measurement, the molecular ion weight (M<sup>+</sup>) of the peptide substrate was used as an internal calibration standard. The mass spectrum of the T=120 min. digested sample is shown in figure 4.

5. MS data analysis and epitope identification:

To assign the measured mass peaks, the computer program MS-Product, a tool from the UCSF Mass Spectrometry Facility (<http://prospector.ucsf.edu/ucsfhtml3.4/msprod.htm>), was used to generate all possible fragments (N- and C-terminal ions, and internal fragments) and their corresponding molecular weights. Due to the sensitivity of the mass spectrometer, average molecular weight was used. The mass peaks observed over the course of the digestion were identified as summarized in Table 4.

10 Fragments co-C-terminal with 8-10 amino acid long sequences predicted to bind HLA by the SYFPEITHI or NIH algorithms were chosen for further study. The digestion and prediction steps of the procedure can be usefully practiced in any order. Although the substrate peptide used in proteasomal digest described here was specifically designed to include predicted HLA-A2.1 binding sequences, the actual products of digestion can be checked after the fact for actual or predicted binding to other MHC molecules. Selected results are shown in Table 5.

Table 4. SSX-2<sub>31-68</sub> Mass Peak Identification.

MS PEAK (measured)	PEPTIDE	SEQUENCE	CALCULATED MASS (MH <sup>+</sup> )
988.23	31-37	YFSKEEW	989.08
1377.68±2.3			
8	31-40	YFSKEEWEKM	1377.68
1662.45±1.3			
0	31-43	YFSKEEWEKMKAS	1663.90
2181.72±0.8			
5	31-47	YFSKEEWEKMKASEKIF	2181.52
2346.6	31-48	YFSKEEWEKMKASEKIFY	2344.71
1472.16±1.5			
4	38-49	EKMASEKIFYV	1473.77
2445.78±1.1			
8	31-49*	YFSKEEWEKMKASEKIFYV	2443.84
2607.	31-50	YFSKEEWEKMKASEKIFYVY	2607.02
1563.3	50-61	YMKRKYEAMTKL	1562.93
3989.9	31-61	YFSKEEWEKMKASEKIFYVYMKRKYEAMTKL	3987.77
1603.74±1.5			
3	51-63	MKRKYEAMTKLGF	1603.98
1766.45±1.5	50-63	YMKRKYEAMTKLGF	1767.16
1866.32±1.2			
2	49-63	VYMKRKYEAMTKLGF	1866.29
4192.6	31-63	YFSKEEWEKMKASEKIFYVYMKRKYEAMTKLG	4192.00
4392.1	31-65**	F	4391.25
		YFSKEEWEKMKASEKIFYVYMKRKYEAMTKLG	
		FKA	

**Boldface** sequence correspond to peptides predicted to bind to MHC.

5

\* On the basis of mass alone this peak could also have been assigned to the peptide 32-50, however proteasomal removal of just the N-terminal amino acid is unlikely. N-terminal sequencing (below) verifies the assignment to 31-49.

\*\* On the basis of mass this fragment might also represent 33-68. N-terminal sequencing below is consistent with the assignment to 31-65.

**Table 5. Predicted HLA binding by proteasomally generated fragments**

SEQ ID NO.	PEPTIDE	HLA	SYFPEITHI	NIH
11	FSKEEWEKM	B*3501	NP†	90
12	KMKASEKIF	B*08	17	<5
13 & (14)	(K) MKASEKIFY	A1	19 (19)	<5
15 & (16)	(M) KASEKIFYV	A*0201	22 (16)	1017
		B*08	17	<5
		B*5101	22 (13)	60
		B*5102	NP	133
		B*5103	NP	121
17 & (18)	(K) ASEKIFYVY	A1	34 (19)	14
19 & (20)	(K) RKYEAMTKL	A*0201	15	<5
		A26	15	NP
		B14	NP	45 (60)
		B*2705	21	15
		B*2709	16	NP
		B*5101	15	<5
21	KYEAMTKLGF	A1	16	<5
22	YEAMTKLGF	A24	NP	300
		B*4403	NP	80
23	EAMTKLGF	B*08	22	<5

†No prediction

5

As seen in Table 5, N-terminal addition of authentic sequence to epitopes can generate epitopes for the same or different MHC restriction elements. Note in particular the pairing of (K)RKYEAAMTKL (SEQ ID NOS 19 and (20)) with HLA-B14, where the 10-mer has a longer predicted halftime of dissociation than the co-C-terminal 9-mer. Also note the case of the 10-mer KYEAMTKLGF (SEQ ID NO. 21) which can be used as a vaccine useful with several MHC types by relying on N-terminal trimming to create the epitopes for HLA-B\*4403 and -B\*08.

10 6. HLA-A0201 binding assay:  
 Binding of the candidate epitope KASEKIFYV, SSX-2<sub>41-49</sub>, (SEQ ID NO. 15) to HLA-A2.1 was assayed using a modification of the method of Stauss et al., (Proc Natl Acad Sci USA 89(17):7871-5 (1992)). Specifically, T2 cells, which express empty or unstable MHC molecules on their surface, were washed twice with Iscove's modified Dulbecco's medium (IMDM) and cultured overnight in serum-free AIM-V medium (Life Technologies, Inc., Rockville, MD) supplemented with human  $\beta$ 2-microglobulin at 3  $\mu$ g/ml (Sigma, St. Louis, MO) and added peptide,

at 800, 400, 200, 100, 50, 25, 12.5, and 6.25  $\mu$ g/ml. in a 96-well flat-bottom plate at  $3 \times 10^5$  cells/200  $\mu$ l/well. Peptide was mixed with the cells by repipeting before distributing to the plate (alternatively peptide can be added to individual wells), and the plate was rocked gently for 2 minutes. Incubation was in a 5% CO<sub>2</sub> incubator at 37°C. The next day the unbound peptide was 5 removed by washing twice with serum free RPMI medium and a saturating amount of anti-class I HLA monoclonal antibody, fluorescein isothiocyanate (FITC)-conjugated anti-HLA A2, A28 (One Lambda, Canoga Park, CA) was added. After incubation for 30 minutes at 4°C, cells were washed 10 3 times with PBS supplemented with 0.5% BSA, 0.05% (w/v) sodium azide, pH 7.4-7.6 (staining buffer). (Alternatively W6/32 (Sigma) can be used as the anti-class I HLA monoclonal antibody the cells washed with staining buffer and then incubated with fluorescein isothiocyanate (FITC)-conjugated goat F(ab') antimouse-IgG (Sigma) for 30 min at 4°C and washed 3 times as before.) The cells were resuspended in 0.5 ml staining buffer. The analysis of surface HLA-A2.1 molecules 15 stabilized by peptide binding was performed by flow cytometry using a FACScan (Becton Dickinson, San Jose, CA). If flow cytometry is not to be performed immediately the cells can be fixed by adding a quarter volume of 2% paraformaldehyde and storing in the dark at 4°C.

The results of the experiment are shown in Figure 5. SSX-2<sub>41-49</sub> (SEQ ID NO. 15) was found to bind HLA-A2.1 to a similar extent as the known A2.1 binder FLPSDYFPSV (HBV<sub>18-27</sub>; SEQ ID NO: 24) used as a positive control. An HLA-B44 binding peptide, AEMGKYSFY (SEQ ID NO: 25), was used as a negative control. The fluorescence obtained from the negative control 20 was similar to the signal obtained when no peptide was used in the assay. Positive and negative control peptides were chosen from Table 18.3.1 in *Current Protocols in Immunology* p. 18.3.2, John Wiley and Sons, New York, 1998.

7. Immunogenicity:

A. In vivo immunization of mice.

HHD1 transgenic A\*0201 mice (Pascolo, S., et al. *J. Exp. Med.* 185:2043-2051, 1997) 25 were anesthetized and injected subcutaneously at the base of the tail, avoiding lateral tail veins, using 100  $\mu$ l containing 100 nmol of SSX-2<sub>41-49</sub> (SEQ ID NO. 15) and 20  $\mu$ g of HTL epitope peptide in PBS emulsified with 50  $\mu$ l of IFA (incomplete Freund's adjuvant).

B. Preparation of stimulating cells (LPS blasts).

Using spleens from 2 naive mice for each group of immunized mice, un-immunized mice 30 were sacrificed and the carcasses were placed in alcohol. Using sterile instruments, the top dermal layer of skin on the mouse's left side (lower mid-section) was cut through, exposing the peritoneum. The peritoneum was saturated with alcohol, and the spleen was aseptically extracted. The spleen was placed in a petri dish with serum-free media. Splenocytes were isolated by using 35 sterile plungers from 3 ml syringes to mash the spleens. Cells were collected in a 50 ml conical tubes in serum-free media, rinsing dish well. Cells were centrifuged (12000 rpm, 7 min) and

washed one time with RPMI. Fresh spleen cells were resuspended to a concentration of  $1 \times 10^6$  cells per ml in RPMI-10%FCS (fetal calf serum). 25g/ml lipopolysaccharide and 7  $\mu\text{g}/\text{ml}$  Dextran Sulfate were added. Cell were incubated for 3 days in T-75 flasks at 37°C, with 5% CO<sub>2</sub>. Splenic blasts were collected in 50 ml tubes pelleted (12000 rpm, 7 min) and resuspended to  $3 \times 10^7/\text{ml}$  in RPMI. The blasts were pulsed with the priming peptide at 50  $\mu\text{g}/\text{ml}$ , RT 4hr. mitomycin C-treated at 25 $\mu\text{g}/\text{ml}$ , 37°C, 20 min and washed three times with DMEM.

5 C. In vitro stimulation.

3 days after LPS stimulation of the blast cells and the same day as peptide loading, the primed mice were sacrificed (at 14 days post immunization) to remove spleens as above.  $3 \times 10^6$  splenocytes were co-cultured with  $1 \times 10^6$  LPS blasts/well in 24-well plates at 37°C, with 5% CO<sub>2</sub> in DMEM media supplemented with 10% FCS,  $5 \times 10^{-5}$  M  $\beta$ -mercaptoethanol, 100 $\mu\text{g}/\text{ml}$  streptomycin and 100 IU/ml penicillin. Cultures were fed 5% (vol/vol) ConA supernatant on day 3 and assayed for cytolytic activity on day 7 in a <sup>51</sup>Cr-release assay.

10 D. Chromium-release assay measuring CTL activity.

15 To assess peptide specific lysis,  $2 \times 10^6$  T2 cells were incubated with 100  $\mu\text{Ci}$  sodium chromate together with 50  $\mu\text{g}/\text{ml}$  peptide at 37°C for 1 hour. During incubation they were gently shaken every 15 minutes. After labeling and loading, cells were washed three times with 10 ml of DMEM-10% FCS, wiping each tube with a fresh Kimwipe after pouring off the supernatant. Target cells were resuspended in DMEM-10% FBS  $1 \times 10^5/\text{ml}$ . Effector cells were adjusted to 20  $1 \times 10^7/\text{ml}$  in DMEM-10% FCS and 100  $\mu\text{l}$  serial 3-fold dilutions of effectors were prepared in U-bottom 96-well plates. 100  $\mu\text{l}$  of target cells were added per well. In order to determine spontaneous release and maximum release, six additional wells containing 100  $\mu\text{l}$  of target cells were prepared for each target. Spontaneous release was revealed by incubating the target cells with 100  $\mu\text{l}$  medium; maximum release was revealed by incubating the target cells with 100 $\mu\text{l}$  of 2% SDS. Plates were then centrifuged for 5 min at 600 rpm and incubated for 4 hours at 37°C in 5% CO<sub>2</sub> and 80% humidity. After the incubation, plates were then centrifuged for 5 min at 1200 rpm. Supernatants were harvested and counted using a gamma counter. Specific lysis was determined as follows: % specific release = [(experimental release - spontaneous release)/(maximum release - spontaneous release)]  $\times$  100.

25 30 Results of the chromium release assay demonstrating specific lysis of peptide pulsed target cells are shown in figure 6.

35 8. Cross-reactivity with other SSX proteins:

SSX-2<sub>41-49</sub> (SEQ ID NO. 15) shares a high degree of sequence identity with the same region of the other SSX proteins. The surrounding regions have also been generally well conserved. Thus the housekeeping proteasome can cleave following V<sub>49</sub> in all five sequences. Moreover, SSX<sub>41-49</sub> is

predicted to bind HLA-A\*0201 (see Table 6). CTL generated by immunization with SSX-2<sub>41-49</sub> cross-react with tumor cells expressing other SSX proteins.

**Table 6. SSX<sub>41-49</sub> – A\*0201 Predicted Binding**

SEQ ID NO.	Family Member	Sequence	SYFPEITHI Score	NIH Score
15	SSX-2	KASEKIFYV	22	1017
26	SSX-1	KYSEKISYV	18	1.7
27	SSX-3	KVSEKIVYV	24	1105
28	SSX-4	KSSEKIVYV	20	82
29	SSX-5	KASEKITYV	22	175

5 **Example 4**

**Cluster Analysis (PSMA<sub>163-192</sub>).**

A peptide, AFSPQGMPEGDLVYVNYARTEDFFKLERDM, PSMA<sub>163-192</sub>, (SEQ ID NO. 30), containing an A1 epitope cluster from prostate specific membrane antigen, PSMA<sub>168-190</sub> (SEQ ID NO. 31) was synthesized using standard solid-phase F-moc chemistry on a 433A ABI Peptide synthesizer. After side chain deprotection and cleavage from the resin, peptide first dissolved in formic acid and then diluted into 30% Acetic acid, was run on a reverse-phase preparative HPLC C4 column at following conditions: linear AB gradient ( 5% B/min) at a flow rate of 4 ml/min, where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile. A fraction at time 16.642 min containing the expected peptide, as judged by mass spectrometry, was pooled and lyophilized. The peptide was then subjected to proteasome digestion and mass spectrum analysis essentially as described above. Prominent peaks from the mass spectra are summarized in Table 7.

**Table 7. PSMA<sub>163-192</sub> Mass Peak Identification.**

PEPTIDE	SEQUENCE	CALCULATE D MASS (MH <sup>+</sup> )
163-177	AFSPQGMPEGDLVYV	1610.0
178-189	NYARTEDFFKLE	1533.68
170-189	PEGDLVYVNYARTEDFFKLE	2406.66
178-191	NYARTEDFFKLERD	1804.95
170-191	PEGDLVYVNYARTEDFFKLERD	2677.93
178-192	NYARTEDFFKLERDM	1936.17
163-176	AFSPQGMPEGDLVY	1511.70
177-192	VNYARTEDFFKLERDM	2035.30
163-179	AFSPQGMPEGDLVYVNY	1888.12

180-192	ARTEDFFKLERDM	1658.89
163-183	<b>AFSPQGMPEGDLVYVNYARTE</b>	2345.61
184-192	DFFKLERDM	1201.40
176-192	YVNYARTEDFFKLERDM	2198.48
167-185	QGMPEGDLVYVNYARTEDF	2205.41
178-186	<b>NYARTEDFF</b>	1163.22

**Boldface** sequences correspond to peptides predicted to bind to MHC, see Table 8.

#### N-terminal Pool Sequence Analysis

One aliquot at one hour of the proteasomal digestion (see Example 3 part 3 above) was subjected to N-terminal amino acid sequence analysis by an ABI 473A Protein Sequencer (Applied Biosystems, Foster City, CA). Determination of the sites and efficiencies of cleavage was based on consideration of the sequence cycle, the repetitive yield of the protein sequencer, and the relative yields of amino acids unique in the analyzed sequence. That is if the unique (in the analyzed sequence) residue X appears only in the nth cycle a cleavage site exists n-1 residues before it in the N-terminal direction. In addition to helping resolve any ambiguity in the assignment of mass to sequences, these data also provide a more reliable indication of the relative yield of the various fragments than does mass spectrometry.

For PSMA<sub>163-192</sub> (SEQ ID NO. 30) this pool sequencing supports a single major cleavage site after V<sub>177</sub> and several minor cleavage sites, particularly one after Y<sub>179</sub>. Reviewing the results presented in figures 7A-C reveals the following:

- S at the 3<sup>rd</sup> cycle indicating presence of the N-terminus of the substrate.
- Q at the 5<sup>th</sup> cycle indicating presence of the N-terminus of the substrate.
- N at the 1<sup>st</sup> cycle indicating cleavage after V<sub>177</sub>.
- N at the 3<sup>rd</sup> cycle indicating cleavage after V<sub>175</sub>. Note the fragment 176-192 in Table 7.
- T at the 5<sup>th</sup> cycle indicating cleavage after V<sub>177</sub>.
- T at the 1<sup>st</sup>-3<sup>rd</sup> cycles, indicating increasingly common cleavages after R<sub>181</sub>, A<sub>180</sub> and Y<sub>179</sub>. Only the last of these correspond to peaks detected by mass spectrometry, 163-179 and 180-192, see Table 7. The absence of the others can indicate that they are on fragments smaller than were examined in the mass spectrum.
- K at the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> cycles indicating cleavages after E<sub>183</sub>, Y<sub>179</sub>, and V<sub>177</sub>, respectively, all of which correspond to fragments observed by mass spectroscopy. See Table 7.
- A at the 1<sup>st</sup> and 3<sup>rd</sup> cycles indicating presence of the N-terminus of the substrate and cleavage after V<sub>177</sub>, respectively.
- P at the 4<sup>th</sup> and 8<sup>th</sup> cycles indicating presence of the N-terminus of the substrate.

G at the 6<sup>th</sup> and 10<sup>th</sup> cycles indicating presence of the N-terminus of the substrate.

M at the 7<sup>th</sup> cycle indicating presence of the N-terminus of the substrate and/or cleavage after F<sub>185</sub>.

M at the 15<sup>th</sup> cycle indicating cleavage after V<sub>177</sub>.

5 The 1<sup>st</sup> cycle can indicate cleavage after D<sub>191</sub>, see Table 7.

R at the 4<sup>th</sup> and 13<sup>th</sup> cycle indicating cleavage after V<sub>177</sub>.

R at the 2<sup>nd</sup> and 11<sup>th</sup> cycle indicating cleavage after Y<sub>179</sub>.

10 V at the 2<sup>nd</sup>, 6<sup>th</sup>, and 13<sup>th</sup> cycle indicating cleavage after V<sub>175</sub>, M<sub>169</sub> and presence of the N-terminus of the substrate, respectively. Note fragments beginning at 176 and 170 in Table 7.

Y at the 1<sup>st</sup>, 2<sup>nd</sup>, and 14<sup>th</sup> cycles indicating cleavage after V<sub>175</sub>, V<sub>177</sub>, and presence of the N-terminus of the substrate, respectively.

15 L at the 11<sup>th</sup> and 12<sup>th</sup> cycles indicating cleavage after V<sub>177</sub>, and presence of the N-terminus of the substrate, respectively, is the interpretation most consistent with the other data. Comparing to the mass spectrometry results we see that L at the 2<sup>nd</sup>, 5<sup>th</sup>, and 9<sup>th</sup> cycles is consistent with cleavage after F<sub>186</sub>, E<sub>183</sub> or M<sub>169</sub>, and Y<sub>179</sub>, respectively. See Table 7.

#### Epitope Identification

20 Fragments co-C-terminal with 8-10 amino acid long sequences predicted to bind HLA by the SYFPEITHI or NIH algorithms were chosen for further analysis. The digestion and prediction steps of the procedure can be usefully practiced in any order. Although the substrate peptide used in proteasomal digest described here was specifically designed to include a predicted HLA-A1 binding sequence, the actual products of digestion can be checked after the fact for actual or predicted binding to other MHC molecules. Selected results are shown in Table 8.

**Table 8. Predicted HLA binding by proteasomally generated fragments**

SEQ ID NO	PEPTIDE	HLA	SYFPEITHI	NIH
32 & (33)	(G) MPEGLVY V	A*0201	17 (27)	(2605)
		B*0702	20	<5
		B*5101	22	314
34 & (35)	(Q) GMPEGDLV Y	A1	24 (26)	<5
		A3	16 (18)	36
		B*2705	17	25
		B*5101	15	NP†
37 & (38)	(P) EGDLVYVN Y	A1	27 (15)	12
		A26	23 (17)	NP
39	LVYVNYARTE	A3	21	<5
40 & (41)	(Y) VNYARTED F	A26	(20)	NP
		B*08	15	<5
		B*2705	12	50
42	NYARTEDFF	A24	NP†	100
		Cw*0401	NP	120
43	YARTEDFF	B*08	16	<5
44	RTEDFFKLE	A1	21	<5
		A26	15	NP

†No prediction

5 **HLA-A\*0201 binding assay:**

HLA-A\*0201 binding studies were preformed with PSMA<sub>168-177</sub>, GMPEGDLVYV, (SEQ ID NO. 33) essentially as described in Example 3 above. As seen in figure 8, this epitope exhibits significant binding at even lower concentrations than the positive control peptides. The Melan-A peptide used as a control in this assay (and throughout this disclosure), ELAGIGILTV, is actually a variant of the natural sequence (EAAGIGILTV) and exhibits a high affinity in this assay.

10

**Example 5****Cluster Analysis (PSMA<sub>281-310</sub>).**

Another peptide, RGIAEAVGLPSIPVHPIGYYDAQKLLEKMG, PSMA<sub>281-310</sub>, (SEQ ID NO. 45), containing an A1 epitope cluster from prostate specific membrane antigen, PSMA<sub>283-307</sub> (SEQ ID NO. 46), was synthesized using standard solid-phase F-moc chemistry on a 433A ABI Peptide synthesizer. After side chain deprotection and cleavage from the resin, peptide in ddH<sub>2</sub>O was run on a reverse-phase preparative HPLC C18 column at following conditions: linear AB gradient (5% B/min) at a flow rate of 4 ml/min, where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile. A fraction at time 17.061 min containing the expected peptide as judged by mass spectrometry, was pooled and lyophilized. The peptide was then subjected to proteasome digestion and mass spectrum analysis essentially as described above. Prominent peaks from the mass spectra are summarized in Table 9.

**Table 9. PSMA<sub>281-310</sub> Mass Peak Identification.**

PEPTIDE	SEQUENCE	CALCULATE D MASS (MH <sup>+</sup> )
281-297	<b>RGIAEAVGLPSIPVHPI*</b>	1727.07
286-297	<b>AVGLPSIPVHPI**</b>	1200.46
287-297	<b>VGLPSIPVHPI</b>	1129.38
288-297	<b>GLPSIPVHPI<sup>†</sup></b>	1030.25
298-310	<b>GYYDAQKLLEKMG<sup>‡</sup></b>	1516.5
298-305	<b>GYYDAQKLS<sup>§</sup></b>	958.05
281-305	<b>RGIAEAVGLPSIPVHPIGYYDAQKL</b>	2666.12
281-307	<b>RGIAEAVGLPSIPVHPIGYYDAQKLLE</b>	2908.39
286-307	<b>AVGLPSIPVHPIGYYDAQKLLE<sup>¶</sup></b>	2381.78
287-307	<b>VGLPSIPVHPIGYYDAQKLLE</b>	2310.70
288-307	<b>GLPSIPVHPIGYYDAQKLLE#</b>	2211.57
281-299	<b>RGIAEAVGLPSIPVHPIGY</b>	1947
286-299	<b>AVGLPSIPVHPIGY</b>	1420.69
287-299	<b>VGLPSIPVHPIGY</b>	1349.61
288-299	<b>GLPSIPVHPIGY</b>	1250.48
287-310	<b>VGLPSIPVHPIGYYDAQKLLEKMG</b>	2627.14
288-310	<b>GLPSIPVHPIGYYDAQKLLEKMG</b>	2528.01

**Boldface** sequences correspond to peptides predicted to bind to MHC, see Table 10.

\*By mass alone this peak could also have been 296-310 or 288-303.

\*\*By mass alone this peak could also have been 298-307. Combination of HPLC and mass spectrometry show that at some later time points this peak is a mixture of both species.

† By mass alone this peak could also have been 289-298.

By mass alone this peak could also have been 281-295 or 294-306.  
§ By mass alone this peak could also have been 297-303.  
¶ By mass alone this peak could also have been 285-306.  
# By mass alone this peak could also have been 288-303.  
5 None of these alternate assignments are supported N-terminal pool sequence analysis.

#### N-terminal Pool Sequence Analysis

One aliquot at one hour of the proteasomal digestion (see Example 3 part 3 above) was subjected to N-terminal amino acid sequence analysis by an ABI 473A Protein Sequencer (Applied 10 Biosystems, Foster City, CA). Determination of the sites and efficiencies of cleavage was based on consideration of the sequence cycle, the repetitive yield of the protein sequencer, and the relative yields of amino acids unique in the analyzed sequence. That is if the unique (in the analyzed 15 sequence) residue X appears only in the nth cycle a cleavage site exists n-1 residues before it in the N-terminal direction. In addition to helping resolve any ambiguity in the assignment of mass to sequences, these data also provide a more reliable indication of the relative yield of the various fragments than does mass spectrometry.

For PSMA<sub>281-310</sub> (SEQ ID NO. 45) this pool sequencing supports two major cleavage sites after V<sub>287</sub> and L<sub>297</sub> among other minor cleavage sites. Reviewing the results presented in Fig. 9 reveals the following:

20 S at the 4<sup>th</sup> and 11<sup>th</sup> cycles indicating cleavage after V<sub>287</sub> and presence of the N-terminus of the substrate, respectively.

H at the 8<sup>th</sup> cycle indicating cleavage after V<sub>287</sub>. The lack of decay in peak height at positions 9 and 10 versus the drop in height present going from 10 to 11 can suggest cleavage after A<sub>286</sub> and E<sub>285</sub> as well, rather than the peaks representing latency in the 25 sequencing reaction.

D at the 2<sup>nd</sup>, 4<sup>th</sup>, and 7<sup>th</sup> cycles indicating cleavages after Y<sub>299</sub>, L<sub>297</sub>, and V<sub>294</sub>, respectively. This last cleavage is not observed in any of the fragments in Table 10 or in the alternate assignments in the notes below.

Q at the 6<sup>th</sup> cycle indicating cleavage after L<sub>297</sub>.

30 M at the 10<sup>th</sup> and 12<sup>th</sup> cycle indicating cleavages after Y<sub>299</sub> and L<sub>297</sub>, respectively.

#### Epitope Identification

35 Fragments co-C-terminal with 8-10 amino acid long sequences predicted to bind HLA by the SYFPEITHI or NIH algorithms were chosen for further study. The digestion and prediction steps of the procedure can be usefully practiced in any order. Although the substrate peptide used in proteasomal digest described here was specifically designed to include a predicted HLA-A1 binding sequence, the actual products of digestion can be checked after the fact for actual or predicted binding to other MHC molecules. Selected results are shown in Table 10.

**Table 10.**  
**Predicted HLA binding by proteasomally generated fragments: PSMA<sub>281-310</sub>**

SEQ ID NO.	PEPTIDE	HLA	SYFPEITHI	NH
47 & (48)	(G) LPSIPVH PI	A*0201	16 (24)	(24)
		B*0702/B7	23	12
		B*5101	24	572
		Cw*0401	NP†	20
49 & (50)	(P) IGYYDAQ KL	A*0201	(16)	<5
		A26	(20)	NP
		B*2705	16	25
		B*2709	15	NP
		B*5101	21	57
		Cw*0301	NP	24
51 & (52)	(P) SIPVHPI GY	A1	21 (27)	<5
		A26	22	NP
		A3	16	<5
		B*5101	16	NP
53	IPVHPIGY			
54	YYDAQKLLE	A1	22	<5

†No prediction

5

As seen in Table 10, N-terminal addition of authentic sequence to epitopes can often generate still useful, even better epitopes, for the same or different MHC restriction elements. Note for example the pairing of (G)LPSIPVHPI with HLA-A\*0201, where the 10-mer can be used as a vaccine useful with several MHC types by relying on N-terminal trimming to create the epitopes for HLA-B7, -B\*5101, and Cw\*0401.

10 HLA-A\*0201 binding assay:

HLA-A\*0201 binding studies were preformed with PSMA<sub>288-297</sub>, GLPSIPVHPI, (SEQ ID NO. 48) essentially as described in Examples 3 and 4 above. As seen in figure 8, this epitope exhibits significant binding at even lower concentrations than the positive control peptides.

**Example 6****Cluster Analysis (PSMA<sub>454-481</sub>).**

Another peptide, SSIEGNYTLRVDCTPLMYSLVHNLTKEL, PSMA<sub>454-481</sub>, (SEQ ID NO. 55) containing an epitope cluster from prostate specific membrane antigen, was synthesized by 5 MPS (purity >95%) and subjected to proteasome digestion and mass spectrum analysis as described above. Prominent peaks from the mass spectra are summarized in Table 11.

**Table 11. PSMA<sub>454-481</sub> Mass Peak Identification.**

MS PEAK (measured)	PEPTIDE	SEQUENCE	CALCULATED MASS (MH <sup>+</sup> )
1238.5	454-464	<b>SSIEGNYTLRV</b>	1239.78
1768.38±0.60	454-469	<b>SSIEGNYTLRVDCTPL</b>	1768.99
1899.8	454-470	<b>SSIEGNYTLRVDCTPLM</b>	1900.19
1097.63±0.91	463-471	<b>RVDCTPLMY</b>	1098.32
2062.87±0.68	454-471*	<b>SSIEGNYTLRVDCTPLMY</b>	2063.36
1153	472-481**	<b>SLVHNLTKE</b>	1154.36
1449.93±1.79	470-481	<b>MYSLVHNLTKE</b>	1448.73

10 **Boldface** sequence correspond to peptides predicted to bind to MHC, see Table 12.

\* On the basis of mass alone this peak could equally well be assigned to the peptide 455-472 however proteasomal removal of just the N-terminal amino acid is considered unlikely. If the issue were important it could be resolved by N-terminal sequencing.

\*\*On the basis of mass this fragment might also represent 455-464.

15

**Epitope Identification**

20 Fragments co-C-terminal with 8-10 amino acid long sequences predicted to bind HLA by the SYFPEITHI or NIH algorithms were chosen for further study. The digestion and prediction steps of the procedure can be usefully practiced in any order. Although the substrate peptide used in proteasomal digest described here was specifically designed to include predicted HLA-A2.1 binding sequences, the actual products of digestion can be checked after the fact for actual or predicted binding to other MHC molecules. Selected results are shown in Table 12.

Table 12. Predicted HLA binding by proteasomally generated fragments

<u>SEQ ID NO</u>	<u>PEPTIDE</u>	<u>HLA</u>	<u>SYFPEITHI</u>	<u>NIH</u>
56 & (57)	(S)IEGNYTLRV	A1	(19)	<5
58	EGNYTLRV	A*0201	16 (22)	<5
		B*5101	15	NP†
59 & (60)	(Y)TLRVDCTPL	A*0201	20 (18)	(5)
		A26	16 (18)	NP
		B7	14	40
		B8	23	<5
		B*2705	12	30
		Cw*0301	NP	(30)
61	LRVDCTPLM	B*2705	20	600
		B*2709	20	NP
62 & (63)	(L)RVDCTPLMY	A1	32 (22)	125 (13.5)
		A3	25	<5
		A26	22	NP
		B*2702	NP	(200)
		B*2705	13 (NP)	(1000)

†No prediction

5 As seen in Table 12, N-terminal addition of authentic sequence to epitopes can often generate still useful, even better epitopes, for the same or different MHC restriction elements. Note for example the pairing of (L)RVDCTPLMY (SEQ ID NOS 62 and (63)) with HLA-B\*2702/5, where the 10-mer has substantial predicted halftimes of dissociation and the co-C-terminal 9-mer does not. Also note the case of SIEGNYTLRV (SEQ ID NO 57) a predicted HLA-A\*0201 epitope which can be  
10 used as a vaccine useful with HLA-B\*5101 by relying on N-terminal trimming to create the epitope.

#### HLA-A\*0201 binding assay

15 HLA-A\*0201 binding studies were preformed, essentially as described in Example 3 above, with PSMA<sub>460-469</sub>, TLRVDCTPL, (SEQ ID NO. 60). As seen in figure 10, this epitope was found to bind HLA-A2.1 to a similar extent as the known A2.1 binder FLPSDYFPSV (HBV<sub>18-27</sub>; SEQ ID NO: 24) used as a positive control. Additionally, PSMA<sub>461-469</sub>, (SEQ ID NO. 59) binds nearly as well.

#### ELISPOT analysis: PSMA<sub>463-471</sub> (SEQ ID NO. 62)

20 The wells of a nitrocellulose-backed microtiter plate were coated with capture antibody by incubating overnight at 4°C using 50 µl/well of 4µg/ml murine anti-human γ-IFN monoclonal

antibody in coating buffer (35 mM sodium bicarbonate, 15 mM sodium carbonate, pH 9.5). Unbound antibody was removed by washing 4 times 5 min. with PBS. Unbound sites on the membrane then were blocked by adding 200 $\mu$ l/well of RPMI medium with 10% serum and incubating 1 hr. at room temperature. Antigen stimulated CD8 $^{+}$  T cells, in 1:3 serial dilutions, 5 were seeded into the wells of the microtiter plate using 100 $\mu$ l/well, starting at 2x10<sup>5</sup> cells/well. (Prior antigen stimulation was essentially as described in Scheibenbogen, C. et al. *Int. J. Cancer* 71:932-936, 1997. PSMA<sub>462-471</sub> (SEQ ID NO. 62) was added to a final concentration of 10 $\mu$ g/ml and IL-2 to 100 U/ml and the cells cultured at 37°C in a 5% CO<sub>2</sub>, water-saturated atmosphere for 10 40 hrs. Following this incubation the plates were washed with 6 times 200  $\mu$ l/well of PBS containing 0.05% Tween-20 (PBS-Tween). Detection antibody, 50 $\mu$ l/well of 2g/ml biotinylated 15 murine anti-human  $\gamma$ -IFN monoclonal antibody in PBS+10% fetal calf serum, was added and the plate incubated at room temperature for 2 hrs. Unbound detection antibody was removed by 20 washing with 4 times 200  $\mu$ l of PBS-Tween. 100 $\mu$ l of avidin-conjugated horseradish peroxidase (Pharmingen, San Diego, CA) was added to each well and incubated at room temperature for 1 hr. Unbound enzyme was removed by washing with 6 times 200  $\mu$ l of PBS-Tween. Substrate was 25 prepared by dissolving a 20 mg tablet of 3-amino 9-ethylcoarbasole in 2.5 ml of N, N-dimethylformamide and adding that solution to 47.5 ml of 0.05 M phosphate-citrate buffer (pH 5.0). 25  $\mu$ l of 30% H<sub>2</sub>O<sub>2</sub> was added to the substrate solution immediately before distributing substrate at 100  $\mu$ l/well and incubating the plate at room temperature. After color development 20 (generally 15-30 min.), the reaction was stopped by washing the plate with water. The plate was air dried and the spots counted using a stereomicroscope.

Figure 11 shows the detection of PSMA<sub>463-471</sub> (SEQ ID NO. 62)-reactive HLA-A1 $^{+}$  CD8 $^{+}$  T cells previously generated in cultures of HLA-A1 $^{+}$  CD8 $^{+}$  T cells with autologous dendritic cells plus the peptide. No reactivity is detected from cultures without peptide (data not shown). In this 25 case it can be seen that the peptide reactive T cells are present in the culture at a frequency between 1 in 2.2x10<sup>4</sup> and 1 in 6.7x10<sup>4</sup>. That this is truly an HLA-A1-restricted response is demonstrated by the ability of anti-HLA-A1 monoclonal antibody to block  $\gamma$ -IFN production; see figure 12.

#### Example 7

##### Cluster Analysis (PSMA<sub>653-687</sub>).

30 Another peptide, FDKNPIVLRMMNDQLMFLERAFIDPLGLPDRPFY PSMA<sub>653-687</sub>, (SEQ ID NO. 64) containing an A2 epitope cluster from prostate specific membrane antigen, PSMA<sub>660-681</sub> (SEQ ID NO 65), was synthesized by MPS (purity >95%) and subjected to proteasome digestion and mass spectrum analysis as described above. Prominent peaks from the mass spectra are summarized in Table 13.

**Table 13. PSMA<sub>653-687</sub> Mass Peak Identification.**

MS PEAK (measured)	PEPTIDE	SEQUENCE	CALCULATED MASS (MH <sup>+</sup> )
906.17±0.65	681-687**	LPDRPFY	908.05
1287.73±0.76	677-687**	DPLGLPDRPFY	1290.47
1400.3±1.79	676-687	IDPLGLPDRPFY	1403.63
1548.0±1.37	675-687	FIDPLGLPDRPFY	1550.80
1619.5±1.51	674-687**	AFIDPLGLPDRPFY	1621.88
1775.48±1.32	673-687*	RAFIDPLGLPDRPFY	1778.07
2440.2±1.3	653-672	FDKSNPIVLRMMNDQLMFLE	2442.93
1904.63±1.56	672-687*	ERAFIGPLGLPDRPFY	1907.19
2310.6±2.5	653-671	FDKSNPIVLRMMNDQLMFL	2313.82
2017.4±1.94	671-687	LERAFIGPLGLPDRPFY	2020.35
2197.43±1.78	653-670	FDKSNPIVLRMMNDQLMF	2200.66

**Boldface** sequence correspond to peptides predicted to bind to MHC, see Table 13.

\* On the basis of mass alone this peak could equally well be assigned to a peptide beginning at 654, however proteasomal removal of just the N-terminal amino acid is considered unlikely. If the issue were important it could be resolved by N-terminal sequencing.

\*\* On the basis of mass alone these peaks could have been assigned to internal fragments, but given the overall pattern of digestion it was considered unlikely.

## Epitope Identification

10 Fragments co-C-terminal with 8-10 amino acid long sequences predicted to bind HLA by the SYFPEITHI or NIH algorithms were chosen for further study. The digestion and prediction steps of the procedure can be usefully practiced in any order. Although the substrate peptide used in proteasomal digest described here was specifically designed to include predicted HLA-A2.1 binding sequences, the actual products of digestion can be checked after the fact for actual or 15 predicted binding to other MHC molecules. Selected results are shown in Table 14.

**Table 14. Predicted HLA binding by proteasomally generated fragments**

SEQ ID NO	PEPTIDE	HLA	SYFPEITHI	NIH
66 & (67)	(R)MMNDQLMFL L	A*0201	24 (23)	1360 (722)
		A*0205	NP†	71 (42)
		A26	15	NP
		B*2705	12	50
68	RMMNDQLMFL	B*2705	17	75

†No prediction

5

As seen in Table 14, N-terminal addition of authentic sequence to epitopes can generate still useful, even better epitopes, for the same or different MHC restriction elements. Note for example the pairing of (R)MMNDQLMFL (SEQ ID NOS. 66 and (67)) with HLA-A\*02, where the 10-mer retains substantial predicted binding potential.

10

#### **HLA-A\*0201 binding assay**

15

HLA-A\*0201 binding studies were preformed, essentially as described in Example 3 above, with PSMA<sub>663-671</sub>, (SEQ ID NO. 66) and PSMA<sub>662-671</sub>, RMMNDQLMFL (SEQ NO. 67). As seen in figures 10, 13 and 14, this epitope exhibits significant binding at even lower concentrations than the positive control peptide (FLPSDYFPSV (HBV<sub>18-27</sub>); SEQ ID NO: 24). Though not run in parallel, comparison to the controls suggests that PSMA<sub>662-671</sub> (which approaches the Melan A peptide in affinity) has the superior binding activity of these two PSMA peptides.

15

#### **Example 8**

20

##### **Vaccinating with epitope vaccines.**

20

###### **1. Vaccination with peptide vaccines:**

###### **A. Intranodal delivery**

25

A formulation containing peptide in aqueous buffer with an antimicrobial agent, an antioxidant, and an immunomodulating cytokine, was injected continuously over several days into the inguinal lymph node using a miniature pumping system developed for insulin delivery (MiniMed; Northridge, CA). This infusion cycle was selected in order to mimic the kinetics of antigen presentation during a natural infection.

25

###### **B. Controlled release**

30

A peptide formulation is delivered using controlled PLGA microspheres as is known in the art, which alter the pharmacokinetics of the peptide and improve immunogenicity. This formulation is injected or taken orally.

5           C.     Gene gun delivery

A peptide formulation is prepared wherein the peptide is adhered to gold microparticles as is known in the art. The particles are delivered in a gene gun, being accelerated at high speed so as to penetrate the skin, carrying the particles into dermal tissues that contain pAPCs.

10           D.     Aerosol delivery

A peptide formulation is inhaled as an aerosol as is known in the art, for uptake into appropriate vascular or lymphatic tissue in the lungs.

15           2.     Vaccination with nucleic acid vaccines:

10           A nucleic acid vaccine is injected into a lymph node using a miniature pumping system, such as the MiniMed insulin pump. A nucleic acid construct formulated in an aqueous buffered solution containing an antimicrobial agent, an antioxidant, and an immunomodulating cytokine, is delivered over a several day infusion cycle in order to mimic the kinetics of antigen presentation during a natural infection.

15           Optionally, the nucleic acid construct is delivered using controlled release substances, such as PLGA microspheres or other biodegradable substances. These substances are injected or taken orally. Nucleic acid vaccines are given using oral delivery, priming the immune response through uptake into GALT tissues. Alternatively, the nucleic acid vaccines are delivered using a gene gun, wherein the nucleic acid vaccine is adhered to minute gold particles. Nucleic acid constructs can also be inhaled as an aerosol, for uptake into appropriate vascular or lymphatic tissue in the lungs.

20           **Example 9**

Assays for the effectiveness of epitope vaccines.

1.     Tetramer analysis:

25           Class I tetramer analysis is used to determine T cell frequency in an animal before and after administration of a housekeeping epitope. Clonal expansion of T cells in response to an epitope indicates that the epitope is presented to T cells by pAPCs. The specific T cell frequency is measured against the housekeeping epitope before and after administration of the epitope to an animal, to determine if the epitope is present on pAPCs. An increase in frequency of T cells specific to the epitope after administration indicates that the epitope was presented on pAPC.

30           2.     Proliferation assay:

35           Approximately 24 hours after vaccination of an animal with housekeeping epitope, pAPCs are harvested from PBMCs, splenocytes, or lymph node cells, using monoclonal antibodies against specific markers present on pAPCs, fixed to magnetic beads for affinity purification. Crude blood or splenocyte preparation is enriched for pAPCs using this technique. The enriched pAPCs are then used in a proliferation assay against a T cell clone that has been generated and is specific for the housekeeping epitope of interest. The pAPCs are coincubated with the T cell clone and the T cells are monitored for proliferation activity by measuring the incorporation of radiolabeled

thymidine by T cells. Proliferation indicates that T cells specific for the housekeeping epitope are being stimulated by that epitope on the pAPCs.

3. Chromium release assay:

5 A human patient, or non-human animal genetically engineered to express human class I MHC, is immunized using a housekeeping epitope. T cells from the immunized subject are used in a standard chromium release assay using human tumor targets or targets engineered to express the same class I MHC. T cell killing of the targets indicates that stimulation of T cells in a patient would be effective at killing a tumor expressing a similar TuAA.

**Example 10**

10 **Induction of CTL response with naked DNA is efficient by Intra-lymph node immunization.**

In order to quantitatively compare the CD8<sup>+</sup> CTL responses induced by different routes of immunization a plasmid DNA vaccine (pEGFPL33A) containing a well-characterized immunodominant CTL epitope from the LCMV-glycoprotein (G) (gp33; amino acids 33-41) (Oehen, S., et al. *Immunology* 99, 163-169 2000) was used, as this system allows a comprehensive assessment of antiviral CTL responses. Groups of 2 C57BL/6 mice were immunized once with titrated doses (200-0.02 $\mu$ g) of pEGFPL33A DNA or of control plasmid pEGFP-N3, administered i.m. (intramuscular), i.d. (intradermal), i.spl. (intrasplenic), or i.ln. (intra-lymph node). Positive control mice received 500 pfu LCMV i.v. (intravenous). Ten days after immunization spleen cells were isolated and gp33-specific CTL activity was determined after secondary *in vitro* restimulation. As shown in Fig. 15, i.m. or i.d. immunization induced weakly detectable CTL responses when high doses of pEGFPL33A DNA (200 $\mu$ g) were administered. In contrast, potent gp33-specific CTL responses were elicited by immunization with only 2 $\mu$ g pEGFPL33A DNA i.spl. and with as little as 0.2 $\mu$ g pEGFPL33A DNA given i.ln. (figure 15; symbols represent individual mice and one of three similar experiments is shown). Immunization with the control pEGFP-N3 DNA did not elicit any detectable gp33-specific CTL responses (data not shown).

**Example 11**

**Intra-lymph node DNA immunization elicits anti-tumor immunity.**

To examine whether the potent CTL responses elicited following i.ln. immunization were able to confer protection against peripheral tumors, groups of 6 C57BL/6 mice were immunized 30 three times at 6-day intervals with 10 $\mu$ g of pEGFPL33A DNA or control pEGFP-N3 DNA. Five days after the last immunization small pieces of solid tumors expressing the gp33 epitope (EL4-33) were transplanted s.c. into both flanks and tumor growth was measured every 3-4d. Although the EL4-33 tumors grew well in mice that had been repetitively immunized with control pEGFP-N3 DNA (figure 16), mice which were immunized with pEGFPL33A DNA i.ln. rapidly eradicated the peripheral EL4-33 tumors (figure 16).

**Example 12**

**Differences in lymph node DNA content mirrors differences in CTL response following intra-lymph node and intramuscular injection.**

5 pEFGPL33A DNA was injected i.ln. or i.m. and plasmid content of the injected or draining lymph node was assessed by real time PCR after 6, 12, 24, 48 hours, and 4 and 30 days. At 6, 12, and 24 hours the plasmid DNA content of the injected lymph nodes was approximately three orders of magnitude greater than that of the draining lymph nodes following i.m. injection. No plasmid DNA was detectable in the draining lymph node at subsequent time points (Fig. 17). This is consonant with the three orders of magnitude greater dose needed using i.m. as compared to i.ln. 10 injections to achieve a similar levels of CTL activity. CD8<sup>+</sup> knockout mice, which do not develop a CTL response to this epitope, were also injected i.ln. showing clearance of DNA from the lymph node is not due to CD8<sup>+</sup> CTL killing of cells in the lymph node. This observation also supports the conclusion that i.ln. administration will not provoke immunopathological damage to the lymph node.

**15 Example 13**

**Administration of a DNA plasmid formulation of a therapeutic vaccine for melanoma to humans.**

SYNCHROTOPE TA2M, a melanoma vaccine, encoding the HLA-A2-restricted tyrosinase epitope SEQ ID NO. 1 and epitope cluster SEQ ID NO. 69, was formulated in 1% Benzyl alcohol, 20 1% ethyl alcohol, 0.5mM EDTA, citrate-phosphate, pH 7.6. Aliquots of 80, 160, and 320 µg DNA/ml were prepared for loading into MINIMED 407C infusion pumps. The catheter of a SILHOUETTE infusion set was placed into an inguinal lymph node visualized by ultrasound imaging. The assembly of pump and infusion set was originally designed for the delivery of insulin to diabetics and the usual 17mm catheter was substituted with a 31mm catheter for this application. 25 The infusion set was kept patent for 4 days (approximately 96 hours) with an infusion rate of about 25 µl/hour resulting in a total infused volume of approximately 2.4 ml. Thus the total administered dose per infusion was approximately 200, and 400 µg; and can be 800 µg, respectively, for the three concentrations described above. Following an infusion subjects were given a 10 day rest period before starting a subsequent infusion. Given the continued residency of plasmid DNA in the 30 lymph node after administration (as in example 12) and the usual kinetics of CTL response following disappearance of antigen, this schedule will be sufficient to maintain the immunologic CTL response.

**Example 14****Additional Epitopes.**

The methodologies described above, and in particular in examples 3-7, have been applied to additional synthetic peptide substrates, leading to the identification of further epitopes as set for 5 the in tables 15-36 below. The substrates used here were designed to identify products of housekeeping proteasomal processing that give rise to HLA-A\*0201 binding epitopes, but additional MHC-binding reactivities can be predicted, as discussed above. Many such reactivities are disclosed, however, these listings are meant to be exemplary, not exhaustive or limiting. As 10 also discussed above, individual components of the analyses can be used in varying combinations and orders. The digests of the NY-ESO-1 substrates 136-163 and 150-177 (SEQ ID NOS. 254 and 255, respectively) yielded fragments that did not fly well in MALDI-TOF mass spectrometry. However, they were quite amenable to N-terminal peptide pool sequencing, thereby allowing 15 identification of cleavage sites. Not all of the substrates necessarily meet the formal definition of an epitope cluster as referenced in example 3. Some clusters are so large, e.g. NY-ESO-1<sub>86-171</sub>, that it was more convenient to use substrates spanning only a portion of this cluster. In other cases, substrates were extended beyond clusters meeting the formal definition to include neighboring 20 predicted epitopes. In some instances, actual binding activity may have dictated what substrate was made, as with for example the MAGE epitopes reported here, where HLA binding activity was determined for a selection of peptides with predicted affinity, before synthetic substrates were designed.

Table 15  
GP100: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion  
†Scores are given from the two binding prediction programs referenced above (see example 3).

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI /NIB)†						Comments
			NO	A*0201	A1	A3	B7	B8	
609-644	630-638*	LPHSSSSHWL	88				20/80	16/≤5	
629-638*	QLPHSSSSHWL	89	21/117						*The digestion of 609-644 and 622-650 have generated the same epitopes.
614-622	IYRRRLMK	90				32/20			
613-622	SLIYRRRLMK	91	14/≤5			29/60			
615-622	IYRRRLMK	92						15/≤5	
622-650	630-638*	LPHSSSSHWL	93			20/80			
629-638*	QLPHSSSSHWL	94	21/117				20/80	16/≤5	

Table 16A  
MAGE-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI / NIH)†					
			A*0201	A1	A3	B7	B8	Other
86-109	95-102	ESLFRAVI	95					16/≤5
	93-102	IILESLFRAVI	96	21/≤5		20/≤5		
	93-101	IILESLFRAV	97	23/≤5				
	92-101	CILESLFRAV	98	23/55				
	92-100	CILESLFRA	99	20/138				
263-292	263-271	EFLWGPRAL	100					
	264-271	FLWGPRAL	101					
	264-273	FLWGPRALAE	102	16/≤5	19/≤5			17/≤5
	265-274	LWGPRALAE	103	16/≤5				
	268-276	PRALAETSY	104	15/≤5				
	267-276	GPRALAETSY	105	15/≤5		<15/≤5		B4403 (NIH 7); B3501 (NIH 120)
	269-277	RAIAETSYV	106	18/20				
	271-279	LAETSYVKV	107	19/≤5				
	270-279	ALAETSYVKV	108	30/427	19/≤5≤5			
	272-280	AETSYVKVL	109	15/≤5				B4403 (NIH 36)
	271-280	LAEETSYVKVL	110	18/≤5		<15/≤5		
	274-282	TSYVKVLEY	111		26/≤5			B4403 (NIH 14)
	273-282	ETSYVKVLEY	112		28/6			A26 (R 31), B4403 (NIH 14)
	278-286	KVILEYVKV	113	26/743		16/≤5		

Table 16B  
MAGE-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI /NIH)†				
			A*0201	A1	A3	B7	B8
168-193	168-177	SYVLTCLGL	114		15/<5	<15/20	
	169-177	YVLYTCLGL	115	20/32			17/<5
	170-177	VLVTCLGL	116				
229-258	240-248	TQDLVQEKY	117		29/<5		
	239-248	LTDQLVQEKY	118		23/<5		
	232-240	YGEPRKLLT	119		24/11		
	243-251	LVQEKYLY	120		21/<5	21/<5	A26 (R 28)
	242-251	DLVQEKYLEY	121		22/<5	19/<5	A26 (R 30)
	230-238	SAYGEPRKL	122	21/<5			
272-297	278-286	KVILEYVVK	123	26/43	16/<5		
	277-286	VKVILEYVVK	124	17/<5			
	276-284	YVKVILEYVI	125	15/<5	15/<5	17/<5	
	274-282	TSYVKVLEY	126		26/<5		
	273-282	ETSYVKVLEY	127		28/6		
	283-291	VIKVSARVR	128		20/<5		
	282-291	YVIKVSARVR	129		24/<5		

†Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

Table 17A

MAGE-2: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI/NIH)†					
			SEQ ID NO	A*0201	A1	A3	B7	B8
107-126	115-122	ELVHFILL	130					18/≤5
	113-122	MVELVHFILL	131		21/≤5			A26 (R 22)
	109-116	ISRKMKVEL	132					17/≤5
	108-116	AISRKMKVEL	133	25/7		19/≤5	16/12	26/≤5
	107-116	AAISRKMKVEL	134	22/≤5			14/36	n.p./16
	112-120	KMVELVHFL	135	27/28/00				
	109-117	ISRKMKVELV	136	16/≤5				
	108-117	AISRKMKVELV	137	24/11				
	116-124	LVHFILLKY	138		23/≤5	19/≤5		A26 (R 26)
	115-124	ELVHFILLKY	139		24/≤5	19/5		A26 (R 29)
	111-119	RKMMVELVHF	140					
	158-166	LQLVFGIEV	141	17/168				
	157-166	YLQLVFGIEV	142	24/1215				
	159-167	QLVFGIEVV	143	25/32		18/≤5		
	158-167	LQLVFGIEVV	144	18/20				
	164-172	TEVVEVVPPI	145	16/≤5				
	163-172	GIEVVEVVPPI	146	22/≤5				
	162-170	FGIEVVEVNV	147	19/≤5				
	154-162	ASEYLQLVF	148	22/68				
	153-162	KASEYLQLVF	149		15/≤5			

†Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

**Table 17B**  
**MAGE-2: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI /NIH)†					Other
			A*0201	A1	A3	B7	B8	
213-233	218-225	EEKIVVEEL	150				22/≤5	
	216-225	APEEKIVVEEL	151	15/≤5		22/72		
	216-223	APEEKIVE	152				18/≤5	
	220-228	KIWEELSML	153	26/804	16/≤5		16/≤5	A26 (R 26)
	219-228	EKJWEELSML	154					A26 (R 22)
	271-278	FLWGPRAL	155				17/≤5	
271-291	271-279	FLWGPRALI	156	25/398	16/7			
	278-286	LIETSYVKV	157	23/≤5				
	277-286	AIJETSYVKV	158	30/427	21/≤5			
	276-284	RALIETSYV	159	18/19				B5101 (20/55)
	279-287	JETSYVKVL	160	15/≤5				
	278-287	LIETSYVKVL	161	22/≤5				A26 (R 22)

†Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

**Table 18**  
**MAGE-3: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI /NIH) <sup>†</sup>				
			SEQ ID NO	A*0201	A1	A3	B7
267-286	271-278	FLWGPRAL	162				17/<5
	270-278	EFLWGPRAL	163				
	271-279	FLWGPRALV	164	27/2655	16/<5		
	276-284	RALVETSYV	165	18/19			
	272-280	LWGPRALVE	166		15/<5		
	271-280	FLWGPRALVE	167	15/<5	22/<5		
	272-281	LWGPRALVET	168	16/<5			

<sup>†</sup>Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

Table 19A  
NY-ESO-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFFEETHI/NTH)†					
			NO	A*0201	A1	B3	B7	B8
81-113	82-90	GPESRLLEF	169	16/11	18/≤5	22/≤5		
	83-91	PESRLLEFY	170	15/≤5				B4403 (NTH 18)
	82-91	GPESRLLEFY	171	25/11				
	84-92	ESRLLEFYL	172			19/8		
	86-94	RLLEFYLAM	173	21/430	21/≤5			
	88-96	LEFYLAMPF	174					B4403 (NTH 60)
	87-96	LLEFYLAMPF	175	<15/45	18/≤5			
	93-102	AMPFATPMEA	176	15/≤5				
101-133	94-102	MPFATPMEA	177			17/≤5		
	115-123	PLPVPGVLL	178	20/≤5	17/≤5	16/≤5	18/≤5	
	114-123	PPLPVPGVLL	179			23/12		16/≤5
	116-123*	LPVPGVLL	180					
	103-112	ELARRSLAQD	181	15/≤5	20/≤5			
	118-126*	VPGVLLKER	182			17/≤5	16/≤5	
	117-126*	PVPGVLLKEF	183			16/≤5		
	116-123*	LPVPGVLL	184					16/≤5
116-145	127-135	TVSGNLTI	185	21/≤5	19/≤5			
	126-135	FTVSGNLTI	186	20/≤5				
	120-128	GVLKKEFTV	187	20/130	18/≤5			
	121-130	VLLKKEFTVSG	188	17/≤5	18/≤5			
	122-130	LIKKEFTVSG	189	20/≤5	18/≤5			
	118-126*	VPGVLLKEF	190			17/≤5	16/≤5	
	117-126*	PVPGVLLKEF	191			16/≤5		

†Scores are given from the two binding prediction programs referenced above (see example 3).

Table 19B  
NY-ESO-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI /NHLH) <sup>†</sup>					
			SEQ ID NO	A*0201	A1	B7	B8	22/≤5
136-163 (SEQ ID NO 254)	139-147	AADHRQLQL	192	17/≤5	17/≤5			A26 (R 25)
	148-156	SISSC1QQL	193	24/7				
	147-156	LSISSCLQQQL	194	18/≤5				
	138-147	TAADHRQLQL	195	18/≤5				
150-177 (SEQ ID NO 255)	161-169	WTIQCFLPV	196	18/84				A26 (R 19)
	157-165	SLIMMWITQC	197	18/42				
	150-158	SSCLQQQLSL	198	15/≤5				
	154-162	QQLLQLLMWI	199	15/50				
	151-159	SCLQQQLSLL	200	18/≤5				
	150-159	SSCLQQQLSLL	201	16/≤5				
	163-171	TQCFLPVFL	202	<15/12				
	162-171	TTQCFLPVFL	203	18/≤5				

<sup>†</sup>Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score

**Table 20**  
**PRAME: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI/NHD) <sup>†</sup>				Other
			NO	A*0201	A1	A3	
211-245	219-227	PMQDKKML	204	16<5			16/n.d. A26 (R 20)
	218-227	MPMQDKKML	205				
411-446	428-436	QHLIGLSNL	206	18<5			<15/240
	427-436	LQHLIGLSNL	207	16/8			
	429-436	HLIGLSNL	208				17<5 B15 (R 21)
	431-439	IGLSNLTHV	209	18/7			
	430-439	LIGLSNLTHV	210	24/37			B*5101 (R 22)

<sup>†</sup>Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

**Table 21**  
**PSA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI/NID) <sup>†</sup>					
			SEQ ID NO	A*0201	A1	A3	B7	B8
42-77	53-61	VLVHPQWVLT	211	22/112		<15/6	17/<5	
	52-61	GVLVHPQWVLT	212	17/21	16/<5	<15/30		A26 (R 18)
	52-60	GVLVHPQWV	213	17/124				
	59-67	WVLTAAHCl	214	15/16				
	54-63	LVHPQWVLT	215	19/<5	20/<5			A26 (R 16)
	53-62	VLVHPQWVLT	216	17/22				
55-95	54-62	LVHPQWVLT	217		17/n.d.			
	66-73	CIRNKS	218				26/20	
	65-73	HCIRNKS	219				<15/16	
	56-64	HPQWVLTAA	220			18/<5		
	63-72	AAHCIRNKS	221	17/<5				

<sup>†</sup>Scores are given from the two binding prediction programs referenced above (see example 3). R indicates a SYFPEITHI score.

**Table 22**  
**pSCA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	SEQ ID NO	HLA Binding Predictions (SYFPEITHI /NIH)†				Other
				A*0201	A1	A3	B7	
93-123*	116-123	LLWGPQQL	222					16<5
	115-123	LLLWGPQQL	223	<15/18				
	114-123	GLLLWGPQQL	224	<15/10				
	99-107	ALQPAAAIL	225	26/9	22/5	<15/12	16<5	A26 (R 19)
	98-107	HALQPAAAIL	226	18<5		<15/12		

\*L123 is the C-terminus of the natural protein.

†Scores are given from the two binding prediction programs referenced above (see example 3).

**Table 23**  
**Tyrosinase: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	HLA Binding Predictions (SYFPEITHI/NIH)†						
			SEQ ID NO	A*0201	A1	A3	B7	B8	Other
128-157	128-137	APEKDKFFAY	227	29/6	15/5				B4403 (NIH 14)
	129-137	PEKDKFFAY	228	18/5					21/5
	130-138	EKDFFAYL	229				15/5		
	131-138	KDKFFAYL	230				20/5		
197-228	205-213	PAFLPWHL	231				15/5		
	204-213	APAFLPWHL	232				23/360		
	207-216	FLPWHLRLFL	1	25/1310					<15/8
	208-216	LPWHLRLFL	9	17/26			20/80		24/16
214-223	FLLRWEQEIQ	233			15/5				
	212-220	RLFLLRWEQ	234			16/5			
	191-211	GSEWRDIDF	235	18/68					
	192-200	SEIWRDIDF	236						16/5
466-484	207-215	FLVHRLFL	8	22/540			<15/6		B4403 (NIH 400)
	473-481	RIWSWLLGA	237	19/13			15/5		
	476-497	SWLLGAAMV	238	18/5					
	477-486	WLLGAAMVGA	239	21/194		18/5			
	478-486	LLGAAMVGA	240	19/19		16/5			

†Scores are given from the two binding prediction programs referenced above (see example 3).

**Table 24**  
**PSMA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	SEQ ID NO	HLA Binding Predictions (SYTPEETHI /NHD) <sup>†</sup>				Other
				A*0201	A1	A3	B7	
1-30	4-12	LLHETDSA V	241	25/485		15/≤5		18/≤5 24/≤5
	13-21	ATARRPRWL	242	18/≤5				
53-80	53-61	TPKHNMKAF	243					A26 (R 30) A26 (R 27)
	64-73	ELKAENTKKF	244				17/≤5	
69-77	69-77	NIKKFLH <sup>1</sup> NF	245					A26 (R 24)
	68-77	ENIKKFLH <sup>1</sup> NF	246					
215-244	220-228	AGAKGVILY	247			25/≤5		A26 (R 22)
	468-477	PLMYSLVHNL	248	22/≤5				
457-489	469-477	LMSVLVHNL	249	27/193		<15/9		A26 (R 22)
	463-471	RYDCTPLMY	250		32/125	25/≤5		
503-533	465-473	DCTPLMYS L	251					A26 (R 22)
	507-515	SGMTPRISK L	252	21/≤5			21/≤5	
	506-515	FSGMTPRISK L	253	17/≤5				

<sup>†</sup>This H was reported as Y in the SWISSPROT database.

<sup>†</sup>Scores are given from the two binding prediction programs referenced above (see example 3).

**Table 25A**  
**MAGE-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	HLA type	SYFPEITHI	NIH
	125-132	KAEMLESV	256	B5101	19	n.a.
	124-132	TKAEMLESV	257	A0201	20	<5
	123-132	VTKAEMLESV	258	A0201	20	<5
Mage-1 119-146	128-136	MLESVIKNY	259	A1	28	45
	127-136	EMLESVIKNY	260	A26	24	n.a.
	125-133	KAEMLESVI	261	A3	17	5
	146-153	KASESQL	262	A1	15	<1.0
	145-153	GKASESQL	263	A26	23	<1.0
	147-155	ASESQLQVF	264	B5101	23	100
Mage-1 143-170	153-161	LVFGIDVKE	265	A24	N.A.	4
				B08	16	<1.0
				B5101	17	N.A.
				B2705	17	1
				B2709	16	N.A.
				A1	22	68
				A26	16	N.A.
				A3	16	<1.0

**Table 25B**  
**MAGE-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
Mage-1 99-125	114-121	LLKYRARE	266	B8	25	<1.0
	106-113	VADLVGFL	267	B8	16	<1.0
				B5101	21	N.A.
	105-113	KVADLVGFL	268	A0201	23	44
				A26	25	N.A.
	107-115	ADLVGFLL	269	A3	16	<5
Mage-1 126-154				B0702	14	20
				B2705	14	30
	106-115	VADLVGFLL	270	A0201	17	<5
	114-123	LLKYRAREPV	271	A0201	20	2

**Table 26**  
**MAGE-3: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NH
Mage-3 267-295	271-278	FLWGPRAL	162	B08	17	<5
	270-278	EFLWGPRAL	163	A26	21	N.A.
	271-279	FLWGPRALV	164	A24	N.A.	30
	278-286	LVETSYVKV	272	B1510	16	N.A.
	277-286	ALVETSYVKV	273	A0201	27	2655
	285-293	KVLHHHMVKI	274	A3	16	2
	276-284	RALVETSYV	165	A0201	19	<1.0
	283-291	YVKVLHHMV	275	A0201	17	<1.0
	275-283	PRALVETSY	276	A1	17	<1.0
	274-283	GPRALVETSY	277	A1	15	<1.0
	278-287	LVETSYVKVL	278	A0201	18	<1.0
	272-281	LWGPRALVET	168	A0201	16	<1.0
	271-280	FLWGPRALVE	167	A3	22	<5

Table 27A  
Fibronectin ED-B Preferred Epitopes Revealed by Honssekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
ED-B 14'-21*	4'-5**	<i>TIIPEVPQL</i> <sup>†</sup>	279	A0201	27	7
				A26	28	N.A.
				A3	17	<5
				B8	15	<5
				B1510	15	N.A.
				B2705	17	10
5'-5**	1-10	<i>DTIIPEVPQL</i> <sup>†</sup>	280	B2709	15	N.A.
				A0201	20	<5
				A26	32	N.A.
		<i>EVPQLTDLSF</i>	281	A26	29	N.A.

\*This substrate contains the 14 amino acids from fibronectin flanking ED-B to the N-terminal side.

\*\*These peptides span the junction between the N-terminus of the ED-B domain and the rest of fibronectin.

<sup>†</sup> The *italicized* lettering indicates sequence outside the ED-B domain.

**Table 27B**  
**Fibronectin ED-B: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
ED-B 8-35	23-30	TPLNSSTI	282	B5101	22	N.A.
	18-25	IGLRWTPL	283	B5101	18	N.A.
	17-25	SIGLRWTPL	284	A0201	20	5
	25-33	LNSASTIGY	285	A26	18	N.A.
	24-33	PLNSSTIGY	286	B08	25	<5
	23-31	TPLNSSTI	287	A1	19	<5
				A26	16	<5
				A1	20	<5
				A26	24	N.A.
				A3	16	<5
				B0702	17	8
				B5101	25	440

Table 27C  
Fibronectin ED-B: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
ED-B 20-49	31-38	IGYRITVV	288	B5101	25	N.A.
	30-38	IGYRITVV	289	A0201	23	15
	29-38	IGYRITVV	290	A3	17	<1.0
	23-30	TPLNSSTI	282	B08	15	<1.0
	25-33	INSSTIGY	285	B5101	15	3
	24-33	PLNSSTIGY	286	A0201	26	9
	31-39	IGYRITVVVA	291	A26	18	N.A.
	30-39	IGYRITVVVA	292	A3	18	<5
	23-31	TPLNSSTII	287	B0702	17	8
				B5101	25	440

**Table 28A**  
**CEA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NH
184-191	SLPVSPRL	293	B08	19	<5	
			A0201	15	<5	
183-191	QSLPVSPRL	294	B1510	15		
			B2705	18	10	
186-193	PVSPRLQL	295	B2709	15		
			B08	18	<5	
185-193	LPVSPRLQL	296	B0702	26	180	
			B08	16	<5	
			B5101	19	130	
			A0201	23	21	
184-193	SLPVSPRLQL	297	A26	18	N.A.	
			A3	18	<5	
185-192	LPVSPRLQ	298	B5101	17	N.A.	
			A0201	21	4	
192-200	QLSNGNRTL	299	A26	16	N.A.	
			A3	19	<5	
			B08	17	<5	
			B1510	15		
191-200	LQLSNGNRTL	300	A0201	16	3	
179-187	WVNNQSLPV	301	A0201	16	28	
186-194	PVSPRLQLS	302	A26	17	N.A.	
			A3	15	<5	

Table 28B  
CEA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HIA type	SYFPEITHI	NH
	362-369	SLPVSPRL	303	B08	19	<1.0
	361-369	QSLPVSPRL	304	A0201	15	<1.0
				B2705	18	10
	364-371	PVSPRLQL	305	B2709	15	
	363-371	LPVSPRLQL	306	B08	18	<1.0
				B0702	26	180
				B08	16	<1.0
	362-371	SLPVSPRLQL	307	B5101	19	130
CEA 354-380				A0201	23	21
				A26	18	N.A.
				A24	N.A.	6
	363-370	LPVSPRLQ	308	A3	18	<5
				B5101	17	N.A.
				A0201	22	4
	370-378	QLSNDNRTL	309	A26	16	N.A.
				A3	17	<1.0
				B08	17	<1.0
	369-378	LQLSNDNRTL	310	A0201	16	3
	357-365	WVNNQSLPV	311	A0201	16	28
	360-368	NQSLPVSPR	312	B2705	14	100

Table 28C  
CEA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID	Binding Prediction		
				HLA type	SYFPEITHI	NTH
539-547	SLPVSPRL	313	B08	19	<5	<5
			A0201	15	<5	<5
			B1510	15	<5	<5
542-549	QSLPVSPRL	314	B2705	18	10	
			B2709	15		
	PVSPRLQL	315	B08	18	<5	
541-549	LPVSPRLQL	316	B0702	26	180	
			B08	16	<1.0	
			B5101	19	130	
540-549	SLPVSPRLQL	317	A26	18	N.A.	
			A3	18	<5	
			A0201	23	21	
541-548	LPVSPRLQ	318	B5101	17	N.A.	
			A0201	24	4	
			A26	16	N.A.	
548-556	QLSNGNRTL	319	A3	19	<1.0	
			B08	17	<1.0	
			B1510	15		
547-556	LQLSNGNRTL	320	A0201	16	3	
535-543	WVNGQSLPV	321	A0201	18	28	
533-541	LWWVNNGQSL	322	A0201	15	<1.0	
					<5	

Table 28D  
CEA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction	
				HLA type	SYFPEITHI
CEA 532-558 (continued)	532-541	YLTWWVNGQSL	323	A0201	25
	538-546	GQSLDPVSPR	324	A26	18
			B2705	17	N.A.
					100

**Table 29A**  
**HER2/NEU: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
	30-37	DMKLRUPA	325	B08	19	8
	28-37	GTDMDKLRUPA	326	A1	23	6
42-49	HLDMLRHL	327	B08	17	<5	
41-49	THLDMLRHL	328	A0201	17	<5	
40-49	ETHLDMLRHL	329	B1510	24	N.A.	
36-43	PASPETHL	330	A26	29	N.A.	
35-43	LPASPETHL	331	B5101	17	N.A.	
Her-2 25-52			A0201	15	<5	
			B5101	20	130	
			B5102	N.A.	100	
	RLPASPETHL	332	A0201	20	21	
			A0201	15	<5	
38-46	SPETHLDML	333	B0702	20	24	
37-46	ASPETHLDML	334	B08	18	<5	
42-50	HLDMLRHL Y	335	B5101	18	110	
41-50	THLDMLRHL Y	336	A0201	18	<5	
			A1	29	25	
			A26	20	N.A.	
			A3	17	4	
			A1	18	<1.0	

**Table 29B**  
**HER2/NEU: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NTH
	719-726	ELRKVKVL	337	B08	24	16
	718-726	TELRKVKVL	338	A0201 B08	16 22	1 <5
	717-726	ETELRKVKVL	339	B5101 A1	16 18	<5 2
	715-723	LKETELRKV	340	A26	28	6
	714-723	ILKETELRKV	341	A0201	17	<5
Her-2 705-732	712-720	MRLKKETEL	342	B5101 A0201 B08 B2705 B2709	15 29 22 27 21	<5 8 <5 2000 N.A.
	711-720	QMRILKKETEL	343	A0201	20	2
	717-725	ETELRKVKV	344	B0702 A1	13 18	40 5
	716-725	KETELRKVKV	345	A0201 A26	16 18	N.A.
	706-714	MPNQAAQMRI	346	B0702 B5101	16 22	8 629
	705-714	AMPNQAAQMRI	347	A0201	18	8
	706-715	MPNQAAQMRL	348	B0702	20	80

Table 29C  
HER2/NEU: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
Her-2 954-982	966-973	RPRFRELV	349	B08	20	24
	965-973	CRPRFRELV	350	B5101	18	N.A.
				B2709	18	
				A26	25	N.A.
	968-976	RFRELVSEF	351	A24	N.A.	32
				A3	15	<5
967-976				B08	16	<5
				B2705	19	
				A26	18	N.A.
				A26	21	N.A.
				A24	N.A.	6
	964-972	ECRPRFREL	353	B0702	15	40
				B8	27	640
				B1510	16	<5



Table 30  
NY-ESO-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPETHI	NIH
NY-ESO-1 51-77	67-75	GAASGLNGC	354	A0201	15	<5
	52-60	RASGPGGCA	355	B0702	15	<5
	64-72	PHGGAASGL	356	B1510	21	N.A.
	63-72	GPHGGAASGL	357	B0702	22	80
	60-69	APRGPHGGAA	358	B0702	23	60

**Table 31A**  
**PRAME: Preferred Epitopes Revealed by Housekeeping Proteasomic Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
	112-119	VRPRRWKL	359	B08	19	
				A26	27	N.A.
				A24	N.A.	5
	111-119	EVRPRRWKL	360	A3	19	N.A.
				B0702	15	(B7) 300.00
				B08	26	160
	113-121	RPRRWKLQV	361	B0702	21	(B7) 40.00
				B5101	19	110
PRAME 103-135	114-122	PRRWKLQVL	362	B08	26	<5
				B2705	23	200
	113-122	RPRRWKLQVL	363	B0702	24	(B7) 800.00
				B8	N.A.	160
				B5101	N.A.	61
				B5102	N.A.	61
				A24	N.A.	10
	116-124	RWKLQVLL	364	B08	22	<5
				B2705	17	3
PRAME 161-187	115-124	RRWKLQVLDL	365	A0201	16	<5
	174-182	PVEVLVDLF	366	A26	25	N.A.

**Table 31B**  
**PRAME: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
PRAME 185-215	199-206	VKRRKKNVL	367	B08	27	8
				A0201	16	<1.0
				A26	20	N.A.
	198-206	KVTKRKKNVL	368	A3	22	<1.0
				B08	30	40
				B2705	16	
	197-206	EKVKRKKNVL	369	A26	15	N.A.
	198-205	KVKRKKNV	370	B08	20	6
	201-208	RKKNVNLRL	371	B08	20	<5
				A0201	15	<1.0
PRAME 216-245	200-208	KRKKNVNLRL	372	A26	15	N.A.
				B0702	15	<1.0
				B08	21	<1.0
				B2705	28	
				B2709	25	
	199-208	VKRRKKNVRL	373	A0201	16	<1.0
	189-196	DELFSYLL	374	B0702	16	4
				B5101	15	N.A.
	205-213	VLRLCCKKL	375	A0201	22	3
				A26	17	N.A.
	204-213	NVRLCCKKL	376	B08	25	8
				A0201	17	7
				A26	19	N.A.

Table 31C  
PRAME: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
PRAME 185-215 (continued)	194-202	YIEKVKRK	377	A0201 A26 A3 B08 B2705	20 18 25 20 17	<1.0 N.A. 68 <1.0
	74-81	QAWPFTCL	378	B5101	17	n.a.
	73-81	VQAWPFTCL	379	A0201 A24 B0702	14 n.a. 16	7 5 6
	72-81	MVQAWPFTCL	380	A26	22	n.a.
	81-88	LPLGVLMK	381	A24 B0702	n.a. 13	7 30
	80-88	CLPLGVLMK	382	B5101	18	n.a.
	79-88	TCLPLGVLMK	383	A0201	17	<1.0
	84-92	GVLMKGGQL	384	A3 A26 B08	27 21 21	120 n.a. 4
	81-89	LPLGVLMKG	385	B5101	20	2
	80-89	CLPLGVLMKG	386	A0201	16	<1.0
	76-85	WPFTICPLGV	387	B0702	18	4

**Table 31D**  
**PRAME: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
PRAME 39-65	51-59	ELFPPLFMA	388	A0201 A26	19 23	18 N.A.
	49-57	PREFPPLF	389	B2705	22	
	48-57	LPREFPPLF	390	B2709	19	
	50-58	RELFPPLFM	391	B0702 B2705	19 16	4
	49-58	PREFPPLFM	392	B2705	15	
				A1	16	<1.0

Table 32  
PSA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NH
	239-246	RPSLYTKV	393	B5101	21	N.A.
	238-246	ERPSLYTKV	394	B2705	15	60
	236-243	LPERPSLY	395	B5101	18	N.A.
PSA 232-258	ALPERPSLY	396	A1	19	<1.0	
			A26	22	N.A.	
			A3	26	6	
			B08	16	<1.0	
			B2705	11	15	
			B2709	19	N.A.	
			A0201	20	<1.0	
			A1	19	<1.0	
			A26	25	N.A.	
			A3	26	60	
241-249	SLYTKVVHY	397	B08	20	<1.0	
			B2705	13	75	
			A1	20	<1.0	
			A26	16	N.A.	
240-249	PSLYTKVVHY	398	B0702	21	4	
			B5101	23	110	
239-247	RPSLYTKVV	399				

Table 33A  
PSMA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	HLA type	Binding Prediction	NIH
PSMA 202-228	211-218	GNKVKNNAQ	400	B08	22	<5
	202-209	IARYGKVF	401	B08	18	<5
	217-225	AQLAGAKGV	402	A0201	16	26
	207-215	KVFRGNKVK	403	A3	32	15
	211-219	GNKVKNNAQL	404	B8	33	80
PSMA 255-282	269-277	TPGYPANEY	405	A1	16	<5
	268-277	LTPGYPANEY	406	A1	21	1
	271-279	GYPANEYAY	407	A1	15	<5
	270-279	PGYYPANEYAY	408	A1	19	<5
	266-274	DPLTPGYPA	409	B0702	21	3
PSMA 483-509	492-500	SLYESWTKK	410	A3	27	150
	491-500	KSLYESWTKK	411	A3	18	150
	486-494	EGFEGKSLY	412	A26	21	N.A.
	485-494	DEGFEGKSLY	413	A1	16	<5
	498-506	TKKSTSPEF	414	B08	17	N.A.

Table 33B  
PSMA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
PSMA 483-509 (continued)	497-506	WTKKSPSPER	415	A26	24	N.A.
	492-501	SLYBSWTKKS	416	A0201	16	<5
				A3	16	<5
	725-732	WGEVKRQI	417	B08	17	<5
	724-732	AWGEVKRQI	418	B5101	17	N.A.
	723-732	KAWGEVKRQI	419	B5101	15	6
	723-730	KAWGEYKR	420	A0201	16	<1.0
	722-730	SKAWGEVKR	421	B2705	15	<5
	731-739	QYYVAAFTV	422	A0201	21	177
				A3	21	<1.0
PSMA 721-749				B5101	15	5
	733-741	YYAAFTVQAA	423	A0201	17	6
				A3	20	<1.0
	725-733	WGEVKRQIV	424	A1	26	11
	727-735	EVKRQIYVA	425	A26	22	N.A.
				A3	18	<1.0
	738-746	TVQAAAETL	426	A26	18	N.A.
	737-746	FTVQAAAETL	427	A0201	19	<1.0
				A26	19	N.A.

**Table 33C**  
**PSMA: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPBITHI	NIH
PSMA 721-749 (continued)	729-737	KRQIYVAAF	428	A26 B2705 B2709	16 24 21	N.A. 3000 N.A.
	721-729	PSKAWGEVK	429	A3	20	<1.0
	723-731	KAWGEVKRQ	430	B5101	16	<1.0
	100-108	WKKEFGQLDSV	431	A0201	16	<5
PSMA 95-122	99-108	QWKEFGQLDSV	432	A0201	17	<5
	102-111	ERGLDSVELA	433	A26	16	N.A.

Table 34A  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	HLA type	SYTFPETHI	NIH
SCP-1 117-143	126-134	ELRQKESKL	434	A0201	20	<5
				A26	26	N.A.
				A3	17	<5
	125-134	AELROKESKL	435	B0702	13	(B7)40.00
				B8	34	320
	133-141	KLOENRKI	436	A0201	16	<5
				A0201	20	61
	298-305	QLEEKTKL	437	B08	28	2
				A0201	16	33
	297-305	NQLEEKTKL	438	B2705	19	200
SCP-1 281-308	288-296	LLEESRDKV	439	A0201	25	15
				B5101	15	3
	287-296	FLLLEESRDKV	440	A0201	27	2378
				A26	21	N.A.
	291-299	ESRDKVNQL	441	B08	29	240
				A26	19	N.A.
	290-299	EESRDKVNQL	442	A1	31	11
				A26	17	N.A.
SCP-1 471-498	475-483	EKEVHDLEY	443			
				A1	21	<1.0
	474-483	REKEVHDLEY	444	A1	26	45
				A26	30	N.A.
				A3	16	<5
	477-485	EVHDLLEYSY	446	A1	15	1

**Table 34B**  
**SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	HLA type	Binding Prediction
SCP-1 471-498 (continued)	477-485	EVHDLEYSY		A26	29
	477-486	EVHDLEYSYC	447	A3	19
	502-509	KLSSKREL	448	A26	<1.0
	508-515	ELKNTEYF	449	B08	22
	507-515	RELKNTEYF	450	B2705	26
	496-503	KRGQRPKL	451	B4403	4
	494-503	LPKRGQRPKL	452	B08	N.A.
	509-517	LKNTEYFTL	453	B0702	N.A.
	508-517	ELKNTEYFTL	454	B5101	120
	506-514	KRELKNTEY	455	B3501	N.A.
SCP-1 493-520	502-510	KLSSKRELK	456	A0201	16
	498-506	GQRPKLSSK	457	A0201	16
	497-506	RGQRPKLSSK	458	B2705	18
	500-508	RPKLSSKRE	459	A3	200
				B08	22
					<1.0
					30000

Table 34C  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 570-596	573-580	LEYVREEL	460	B08	19	<5
	572-580	ELEYVREEL	461	A0201 A26 A24	17 23 N.A. N.A.	<1.0 N.A. 9
	571-580	N ELEYVREEL	462	A0201 B08	16 20	N.A. N.A.
	579-587	ELKQKRDEV	463	A0201 A26 B08	19 18 29	<1.0 N.A. 48
	575-583	YVREELKQK	464	A26 A3	17 27	N.A. 2
	632-640	QLNVYERKV	465	A0201	24	70
SCP-1 618-645	630-638	SKQLNVYH	466	A0201	17	<5
	628-636	AESKQLNVY	467	A1	19	<5
	627-636	TAESKQLNVY	468	A26 A1	16 26	N.A. 45

Table 34D  
SCP-1; Preferred Epitopes Revealed by Houskeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 633-660	638-645	IKVNKLEL	469	B08	21	<1.0
	637-645	EIKVNKLEL	470	A0201	17	<1.0
				A26	26	N.A.
				B08	28	8
	636-645	YEKVNKLEL	471	B1510	15	N.A.
	642-650	KIELELESA	472	A0201	17	2
				A3	16	<1.0
	635-643	VYEIKVNKL	473	A0201	18	<1.0
				A24	N.A.	396
				B08	22	<1.0
SCP-1 640-668	634-643	NVYEIKVNKL	474	A0201	24	56
				A26	25	N.A.
				A24	N.A.	6
				A3	15	<5
				B0702	11	(B7) 20
SCP-1 640-668	646-654	ELESAKQKF	475	B08	N.A.	6
	642-650	KLELELESA	476	A26	27	N.A.
				A0201	20	1
	646-654	ELESAKQKF	477	A3	16	<1.0
SCP-1 640-668				A26	27	N.A.

Table 34E  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NH
SCP-1 768-796	771-778	KEKLKREA	478	B08	21	<5
	777-785	EAKENTATL	479	A0201	18	<5
				A26	18	N.A.
				A24	N.A.	5
				B0702	13	12
				B08	28	48
				B5101	20	121
				A0201	16	<5
				A3	17	<5
SCP-1 92-125	776-785	REAKENTATL	480	B5101	17	N.A.
	773-782	KLKREAKENT	481	A0201	23	32
	112-119	EAKIKKKW	482	A26	22	N.A.
	101-109	GLSRVYVSKL	483	A24	N.A.	6
				A3	17	3
				B08	17	<1.0
				A26	21	N.A.
				A24	N.A.	9
				A0201	22	57
				A3	20	9
				B5101	18	5
				A1	31	68
98-106	97-106	NSEGLSRVY	486	A26	18	N.A.
				A1	22	<1.0
102-110		LSRVYVSKLY	488			

Table 34F  
SCP-1; Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPETHI	NIH
SCP-1 92-125 (continued)	101-110	GLSRVYSSKLY	489	A1 A26 A3	18 18 19	<1.0 N.A. 18
	96-105	LENSEGLSRV	490	A0201	17	5
	108-117	KLYKEAEKJK	491	A3	27	150
	949-956	REDRWAVI	492	B5101	15	N.A.
	948-956	MREDRWAVI	493	B2705 B2709	18 18	600 N.A.
	947-956	KMREDRWAVI	494	B5101	15	1
	947-955	KMREDRWAV	495	A0201	21	6
	934-942	TTPGSILKF	496	B08	N.A.	15
	933-942	LTTPGSILKF	497	A0201	22	411
	937-945	GSTLKFQGAI	498	B08	25	N.A.
SCP-1 931-958	945-953	IRKMREDRW	499	B08	23	N.A.
	236-243	RLEMHFKL	500	B08	19	1
	235-243	SRLEMHFKL	501	B08	19	<5
	242-250	KLKEDYEKI	502	A0201	22	4

Table 34G  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPERITI	NTH
SCP-1 232-259 (continued)				A26	16	N.A.
				A3	15	3
				B08	24	<5
				B5101	14	2
	249-257	KIQHLFQEY	503	A1	15	<5
				A26	23	N.A.
SCP-1 310-340				A3	17	<5
	248-257	EKIQHLFQEY	504	A1	15	<5
	233-242	ENSRLEMHF	505	A26	21	N.A.
	236-245	RLEMHIFKLKE	506	A26	19	N.A.
	324-331	LEDIKVSL	507	A1	19	<5
				A3	17	<5
SCP-1 310-340				B08	20	<1.0
				A0201	21	<1.0
				A26	25	N.A.
				A24	N.A.	10
				A3	17	<1.0
				B08	19	<1.0
				B1510	16	N.A.
322-331	KELEDIKVSL	509	A0201		19	22
320-327	LTKELEDI	500	B08		18	<5
319-327	HLTKELEDI	511	A0201		21	<1.0
330-338	SLQRSVSTQ	512	A0201		18	<1.0

Table 34H  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 310-340 (continued)	321-329	TKELEDIKV	513	A1	16	<1.0
	320-329	LTKLELEDKV	514	A0201	19	<1.0
	326-335	DIKVSLQRSV	515	A26	18	N.A.
	281-288	KMKDLTFL	516	B08	20	3
	280-288	NKMKDLTFL	517	A0201	15	1
	279-288	ENKMKDLTFL	518	A26	19	N.A.
	288-296	LLEESRDKV	519	A0201	25	15
	287-296	FLLEESRDKV	520	A0201	15	3
	291-299	ESRDKVNQL	521	A26	21	N.A.
	290-299	EESRDKVNQL	522	B08	29	240
SCP-1 272-305	277-285	EKENKMKDL	523	A26	19	N.A.
	276-285	TEKENKMKDL	524	A26	19	N.A.
	279-287	ENKMKDLTF	525	B08	18	N.A.
	218-225	IEKMITAF	526	B08	28	4
	217-225	NIEKMITAF	527	A26	17	<5
	216-225	SNIEKMITAF	528	A26	26	N.A.
	223-230	TAFEELRV	529	B5101	23	N.A.
	222-230	ITAFEELRV	530	A0201	18	2
	221-230	MTTAFEELRV	531	A0201	18	16

Table 34I  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 211-239 (continued)	220-228	KMITA FEEL	532	A0201	23	50
	219-228	EKMITA FEEL	533	A26	15	N.A.
	227-235	ELRVQAENS	534	A24	N.A.	16
	213-222	DLNSNEKMI	535	A26	19	N.A.
	837-844	WTSAKNTL	536	B08	16	<1.0
SCP-1 836-863	846-854	TPLRKAYTV	537	A0201	15	<1.0
	845-854	STPLPKAYTV	538	B0702	17	<1.0
	844-852	LSIPLIJKAY	539	B08	16	N.A.
	843-852	TLSTPLPKAY	540	B5101	20	4
	842-850	NTLSTPLPK	541	A0201	18	2
	841-850	KNTLSTPLPK	542	A1	23	8
				A26	19	N.A.
				A3	18	2
				A3	16	3
				A3	18	<1.0

**Table 34I**  
**SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 819-845	828-835	ISKDKRDY	543	B08	21	3
	826-835	HGISKDKRDY	544	A26	21	N.A.
	832-840	KRDYILWTSA	545	A1	15	<5
	829-838	SKDKRDYLWT	546	B2705	16	600
	279-286	ENKMKDLT	547	B08	18	<5
	260-268	EINDKEKQV	548	A0201	22	8
SCP-1 260-288	274-282	QITEKENKM	549	A26	17	3
	269-277	SLLLQITTE	550	B08	19	N.A.
	453-460	FEKIAEEL	551	A0201	17	<5
	452-460	QFEKIAEEL	552	B08	17	3
	451-460	KQFEKIAEEL	553	A0201	22	N.A.
	449-456	DNKQFEKJ	554	B08	16	<5
SCP-1 437-464	448-456	YDNKQFEKJ	555	B5101	16	1
	447-456	LYDNKQFEKJ	556	A1	15	<1.0

**Table 34K**  
**SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NH
SCP-1 437-464 (continued)	440-447	LGEKETILL	557	B5101	16	N.A.
	439-447	VLGEKETILL	558	A0201	24	149
				A26	19	N.A.
				B08	29	12
	438-447	KVLGEKETILL	559	A0201	19	24
				A26	20	N.A.
				A24	N.A.	12
				A3	18	<1.0
				B0702	14	20
				A0201	22	3
SCP-1 383-412	390-398	LLRTEQQRL	560	A26	18	N.A.
				B08	22	1.6
				B2705	15	30
				A0201	19	6
	389-398	ELLRTEQQQL	561	A26	24	N.A.
				A3	15	<1.0
	393-401	TEQQRLNEY	562	A1	15	<5
				A26	16	N.A.
	392-401	RTEQQRLNEY	563	A1	31	113
	402-410	EDQLIILTM	564	A26	26	N.A.
	397-406	RLENYEDQLI	565	A0201	18	N.A.
				A3	17	<1.0
					15	<1.0

Table 34K  
SCP-1: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SCP-1 366-394	368-375	KARAAGHSF	566	B08	16	<1.0
	376-384	VVTEFFETTV	567	A0201	19	161
	375-384	FVVTEFFETTV	568	A3	16	<1.0
	377-385	VTEFFETTV	569	A0201	17	106
SCP-1 331-357	376-385	VVTEFFETTV	570	A1	18	2
				A3	16	<5
	344-352	DLQIATNTI	571	A3	22	<5
				A0201	15	<1.0
SCP-1 331-357	347-355	IATNTICQL	572	B5101	17	11
				A0201	19	1
	346-355	QIATNTICQL	573	B08	16	<1.0
				B5101	20	79
				A0201	24	7
				A26	24	N.A.

Table 35  
SSX-4: Preferred Epitopes Recycled by Housekeeping Proteasome Digestion

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
SSX4 45-76	57-65	VM <del>T</del> KLGFKV	574	A0201	21	495
	53-61	LNYEV <del>M</del> TKL	575	A0201	17	7
				A0201	23	172
				A26	21	N.A.
SSX4 52-61	52-61	KLN <del>YEV</del> MTKL	576	A24	N.A.	18
				A3	14	4
				B7	N.A.	4
	66-74	TLPPFM <del>R</del> SK	577	A26	16	N.A.
				A3	25	14
				A0201	15	<5
	110-1118	KIMPKKKPAE	578	A26	15	N.A.
SSX4 98-124	103-112	SLQRIFPKIM	579	A3	16	<5
				A0201	15	8
				A26	16	N.A.
				A3	15	<5

**Table 36**  
**Tyrosinase: Preferred Epitopes Revealed by Housekeeping Proteasome Digestion**

Substrate	Epitope	Sequence	Seq. ID No.	Binding Prediction		
				HLA type	SYFPEITHI	NIH
Tyr 445-474	463-471	YIKSYLEQA	580	A0201	18	<5
	459-467	SFQDYIKSY	581	A26	17	N.A.
458-467	DSFQDYIKSY	582	A1	18	<5	N.A.
	507-514	LPEEKQPL	583	B08	28	5
506-514	QLPEEKQPL	584	B5101	18	N.A.	N.A.
	505-514	KQLPEEKQPL	585	A0201	22	88
Tyr 490-518	507-515	LPEEKQPLL	586	A26	20	N.A.
	506-515	QLPEEKQPLL	587	A24	N.A.	9
497-505	SLLCRHKRK	588	B08	18	<5	N.A.
			A0201	15	28	N.A.
			A24	N.A.	17	
			A0201	15	<5	
			B0702	21	24	
			B08	28	5	
			B5101	21	157	
			A0201	23	88	
			A26	20	N.A.	
			A24	N.A.	7	
			A3	25	15	

**Example 15**  
**Evaluating Likelihood of Epitope Cross-reactivity on Non-target Tissues.**

As noted above PSA is a member of the kallikrein family of proteases, which is itself a subset of the serine protease family. While the members of this family sharing the greatest degree of sequence identity with PSA also share similar expression profiles, it remains possible that individual epitope sequences might be shared with proteins having distinctly different expression profiles. A first step in evaluating the likelihood of undesirable cross-reactivity is the identification of shared sequences. One way to accomplish this is to conduct a BLAST search of an epitope sequence against the SWISSPROT or Entrez non-redundant peptide sequence databases using the "Search for short nearly exact matches" option; hypertext transfer protocol accessible on the world wide web (<http://www>) at "[ncbi.nlm.nih.gov/blast/index.html](http://ncbi.nlm.nih.gov/blast/index.html)". Thus searching SEQ ID NO. 214, WVLTAACI, against SWISSPROT (limited to entries for homo sapiens) one finds four exact matches, including PSA. The other three are from kallikrein 1 (tissue kallikrein), and elastase 2A and 2B. While these nine amino acid segments are identical, the flanking sequences are quite distinct, particularly on the C-terminal side, suggesting that processing may proceed differently and that thus the same epitope may not be liberated from these other proteins. (Please note that kallikrein naming is confused. Thus the kallikrein 1 [accession number P06870] is a different protein than the one [accession number AAD13817] mentioned in the paragraph on PSA above in the section on tumor-associated antigens).

It is possible to test this possibility in several ways. Synthetic peptides containing the epitope sequence embedded in the context of each of these proteins can be subjected to *in vitro* proteasomal digestion and analysis as described above. Alternatively, cells expressing these other proteins, whether by natural or recombinant expression, can be used as targets in a cytotoxicity (or similar) assay using CD8<sup>+</sup> T cells that recognize the epitope, in order to determine if the epitope is processed and presented.

**Example 16**

**Epitope Clusters.**

Known and predicted epitopes are generally not evenly distributed across the sequences of protein antigens. As referred to above, we have defined segments of sequence containing a higher than average density of (known or predicted) epitopes as epitope clusters. Among the uses of epitope clusters is the incorporation of their sequence into substrate peptides used in proteasomal digestion analysis as described herein. Epitope clusters can also be useful as vaccine components. A fuller discussion of the definition and uses of epitope clusters is found in U.S. Patent Application No. 09/561,571 entitled EPITOPE CLUSTERS, previously incorporated by reference.

The following tables (37-60) present 9-mer epitopes predicted for HLA-A2 binding using both the SYFPEITHI and NIH algorithms and the epitope density of regions of overlapping

5 epitopes, and of epitopes in the whole protein, and the ratio of these two densities. (The ratio must exceed one for there to be a cluster by the above definition; requiring higher values of this ratio reflect preferred embodiments). Individual 9-mers are ranked by score and identified by the position of their first amino in the complete protein sequence. Each potential cluster from a protein is numbered. The range of amino acid positions within the complete sequence that the cluster covers is indicated as are the rankings of the individual predicted epitopes it is made up of.

10

**Table 37**  
**BIMAS-NIH/Parker algorithm Results for gp100**

Rank	Start	Score	Rank	Start	Score
1	619	1493	21	416	19
2	602	413	22	25	18
3	162	226	23	566	17
4	18	118	24	603	15
5	178	118	25	384	14
6	273	117	26	13	14
7	601	81	27	290	12
8	243	63	28	637	10
9	606	60	29	639	9
10	373	50	30	485	9
11	544	36	31	453	8
12	291	29	32	102	8
13	592	29	33	399	8
14	268	29	34	456	7
15	47	27	35	113	7
16	585	26	36	622	7
17	576	21	37	69	7
18	465	21	38	604	6
19	570	20	39	350	6
20	9	19	40	583	5

**Table 38**  
**SYFPEITHI (Rammensee algorithm) Results for gp100**

Rank	Start	Score	Rank	Start	Score	Rank	Start	Score
1	606	30	37	291	20	73	60	18
2	162	29	38	269	20	74	17	18
3	456	28	39	2	20	75	613	17
4	18	28	40	610	19	76	599	17
5	602	27	41	594	19	77	572	17
6	598	27	42	591	19	78	557	17
7	601	26	43	583	19	79	556	17
8	597	26	44	570	19	80	512	17
9	13	26	45	488	19	81	406	17
10	585	25	46	446	19	82	324	17
11	449	25	47	322	19	83	290	17
12	4	25	48	267	19	84	101	17
13	603	24	49	250	19	85	95	17
14	576	24	50	205	19	86	635	16
15	453	24	51	180	19	87	588	16
16	178	24	52	169	19	88	584	16
17	171	24	53	88	19	89	577	16
18	11	24	54	47	19	90	559	16
19	619	23	55	10	19	91	539	16
20	280	23	56	648	18	92	494	16
21	268	23	57	605	18	93	482	16
22	592	22	58	604	18	94	468	16
23	544	22	59	595	18	95	442	16
24	465	22	60	571	18	96	413	16
25	399	22	61	569	18	97	408	16
26	373	22	62	450	18	98	402	16
27	273	22	63	409	18	99	286	16
28	243	22	64	400	18	100	234	16
29	566	21	65	371	18	101	217	16
30	563	21	66	343	18	102	211	16
31	485	21	67	298	18	103	176	16
32	384	21	68	209	18	104	107	16
33	350	21	69	102	18	105	96	16
34	9	21	70	97	18	106	80	16
35	463	20	71	76	18	107	16	16
36	397	20	72	69	18	108	14	16
						109	7	16

**Table 39****Prediction of clusters for gp100**

Total AAs: 661

Total 9-mers: 653

SYFPEITHI 16: 109 9-mers

NIH 5: 40 9-mers

	Cluster #	AAs	Epitopes (by Rank)	Epitopes/AA		
				Cluster	Whole Pr	Ratio
SYFPEITHI	1	2 to 26	39, 12, 109, 34, 55, 11, 9, 108, 107, 74, 4	0.440	0.165	2.668
	2	69-115	72, 71, 106, 53, 85, 105, 70, 84, 69, 104	0.213	0.165	1.290
	3	95-115	85, 105, 70, 84, 69	0.238	0.165	1.444
	4	162-188	2, 52, 17, 103, 16, 51	0.222	0.165	1.348
	5	205-225	50, 68, 102, 101	0.190	0.165	1.155
	6	243-258	28, 49	0.125	0.165	0.758
	7	267-306	48, 21, 38, 27, 20, 99, 83, 37, 67	0.225	0.165	1.364
	8	322-332	47, 82	0.182	0.165	1.103
	9	343-358	66, 33	0.125	0.165	0.758
	10	371-381	65, 26	0.182	0.165	1.103
	11	397-421	36, 25, 64, 98, 81, 97, 63, 96	0.320	0.165	1.941
	12	442-476	95, 46, 11, 62, 15, 3, 35, 24, 94	0.257	0.165	1.559
	13	482-502	93, 31, 45, 93	0.190	0.165	1.155
	14	539-552	91, 23	0.143	0.165	0.866
NIH	15	556-627	79, 78, 90, 30, 29, 61, 44, 60, 77, 14, 89, 43, 88, 10, 87, 42, 22, 41, 59, 8, 6, 76, 7, 5, 13, 58, 57, 1, 40, 75, 19	0.431	0.165	2.611
	1	9 to 33	20, 26, 4, 22	0.160	0.061	2.644
	2	268-281	14, 6	0.143	0.061	2.361
	3	290-299	27, 12	0.200	0.061	3.305
	4*	102-121	32, 35	0.100	0.061	1.653
	5*	373-392	10, 25	0.100	0.061	1.653
	6	453-473	31, 34, 18	0.143	0.061	2.361
	7	566-600	23, 19, 17, 40, 16, 13	0.171	0.061	2.833
	8	601-614	7, 2, 24, 38, 9	0.357	0.061	5.902
	9	619-630	1, 36	0.17	0.061	2.754
	10	637-647	28, 29	0.18	0.061	3.005

\*Nearby but not overlapping epitopes

Table 40  
BIMAS-NIH/Parker algorithm Results for PSMA

Rank	Start	Score
1	663	1360
2	711	1055
3	4	485
4	27	400
5	26	375
6	668	261
7	707	251
8	469	193
9	731	177
10	35	67
11	33	64
12	554	59
13	427	50
14	115	47
15	20	40
16	217	26
17	583	24
18	415	19
19	193	14
20	240	12
21	627	11
22	260	10
23	130	10
24	741	9
25	3	9
26	733	8
27	726	7
28	286	6
29	174	5
30	700	5

**Table 41**  
**SYFPEITHI (Rammensee algorithm) Results for PSMA**

Rank	Start	Score	Rank	Start	Score	Rank	Start	Score
1	469	27	31	26	20	61	305	17
2	27	27	32	3	20	62	304	17
3	741	26	33	583	19	63	286	17
4	711	26	34	579	19	64	282	17
5	354	25	35	554	19	65	169	17
6	4	25	36	550	19	66	142	17
7	663	24	37	547	19	67	122	17
8	130	24	38	390	19	68	738	16
9	57	24	39	219	19	69	634	16
10	707	23	40	193	19	70	631	16
11	260	23	41	700	18	71	515	16
12	20	23	42	472	18	72	456	16
13	603	22	43	364	18	73	440	16
14	218	22	44	317	18	74	385	16
15	109	22	45	253	18	75	373	16
16	731	21	46	91	18	76	365	16
17	668	21	47	61	18	77	361	16
18	660	21	48	13	18	78	289	16
19	507	21	49	733	17	79	278	16
20	454	21	50	673	17	80	258	16
21	427	21	51	671	17	81	247	16
22	358	21	52	642	17	82	217	16
23	284	21	53	571	17	83	107	16
24	115	21	54	492	17	84	100	16
25	33	21	55	442	17	85	75	16
26	606	20	56	441	17	86	37	16
27	568	20	57	397	17	87	30	16
28	473	20	58	391	17	88	21	16
29	461	20	59	357	17			
30	200	20	60	344	17			

Table 42

## Prediction of clusters for prostate-specific membrane antigen (PSMA)

Total AAs: 750  
 Total 9-mers: 742  
 SYFPEITHI 16: 88 9-mers  
 NIH 5: 30 9-mers

	Cluster #	Aas	Epitopes (by rank)	Epitopes/AA		
				Cluster	Whole Pr	Ratio
SYFPEITHI	1	3 to 12	32, 6	0.200	0.117	1.705
	2	13-45	13, 12, 88, 31, 2, 87, 25, 86	0.242	0.117	2.066
	3	57-69	9, 47	0.154	0.117	1.311
	4	100-138	84, 83, 15, 24, 67, 8	0.154	0.117	1.311
	5	193-208	40, 30	0.111	0.117	0.947
	6	217-227	82, 14, 39	0.273	0.117	2.324
	7	247-268	81, 45, 80, 11	0.182	0.117	1.550
	8	278-297	79, 64, 23, 63, 78	0.250	0.117	2.131
	9	354-381	5, 59, 22, 77, 43, 76, 75	0.250	0.117	2.131
	10	385-405	74, 38, 58, 57	0.190	0.117	1.623
	11	440-450	73, 56, 55	0.273	0.117	2.324
	12	454-481	20, 72, 29, 1, 42, 28	0.214	0.117	1.826
	13	507-523	17, 71	0.118	0.117	1.003
	14	547-562	37, 36, 35	0.188	0.117	1.598
	15	568-591	27, 53, 34, 33	0.167	0.117	1.420
	16	603-614	13, 26	0.167	0.117	1.420
	17	631-650	70, 69, 52	0.150	0.117	1.278
	18	660-681	18, 7, 17, 51, 50	0.227	0.117	1.937
	19	700-719	41, 10, 4	0.150	0.117	1.278
	20	731-749	16, 49, 68, 3	0.211	0.117	1.794
NIH	1	3 to 12	25, 3	0.200	0.040	5.000
	2	20-43	15, 5, 4, 11, 10	0.208	0.040	5.208
	3*	415-435	18, 13	0.095	0.040	2.381
	4	663-676	1, 6	0.143	0.040	3.571
	5	700-715	30, 7, 3	0.188	0.040	4.688
	6	726-749	27, 9, 26, 24	0.167	0.040	4.167

\*Nearby but not overlapping epitopes

**Table 43**  
**BIMAS-NIH/Parker algorithm Results for PSA**

Rank	Start	Score
1	7	607
2	170	243
3	52	124
4	53	112
5	195	101
6	165	23
7	72	18
8	245	18
9	2	16
10	59	16
11	122	15
12	125	15
13	191	13
14	9	8
15	14	6
16	175	5
17	130	5

Table 44  
SYFPEITHI (Rammensee algorithm) Results for PSA

Rank	Start	Score
1	72	26
2	170	22
3	53	22
4	7	22
5	234	21
6	166	21
7	140	21
8	66	21
9	241	20
10	175	20
11	12	20
12	41	19
13	20	19
14	14	19
15	130	18
16	124	18
17	121	18
18	47	18
19	17	18
20	218	17
21	133	17
22	125	17
23	122	17
24	118	17
25	110	17
26	67	17
27	52	17
28	21	17
29	16	17
30	2	17
31	184	16
32	179	16
33	158	16
34	79	16
35	73	16
36	4	16

Table 45

## Prediction of clusters for prostate specific antigen (PSA)

Total AAs: 261

Total 9-mers: 253

SYFPEITHI 16: 36 9-mers

NIH 5: 17 9-mers

				Epitopes/AA			
		Cluster #	AAs	Epitopes (by rank)	Cluster	Whole Pr	Ratio
SYFPEITHI	1	2 to 29	30, 36, 4, 11, 14, 29, 19, 13, 28	0.321	0.138	2.330	
	2	41-61	12, 18, 27, 3	0.190	0.138	1.381	
	3	66-87	8, 26, 1, 35, 34	0.227	0.138	1.648	
	4	110-148	25, 24, 17, 23, 16, 22, 15, 21, 7	0.184	0.138	1.332	
	5	158-192	33, 6, 2, 10, 32, 31	0.171	0.138	1.243	
	6	234-249	5, 9	0.125	0.138	0.906	
	7*	118-133	24, 17, 23, 16, 22	0.313	0.138	2.266	
	8*	118-138	24, 17, 23, 16, 22, 15	0.286	0.138	2.071	
NIH	1	2-22	9, 1, 14, 15	0.190	0.065	2.924	
	2	52-67	3, 4, 10	0.188	0.065	2.879	
	3	122-138	11, 12, 17	0.176	0.065	2.709	
	4	165-183	6, 2, 16	0.158	0.065	2.424	
	5	191-203	13, 5	0.154	0.065	2.362	
	6**	52-80	3, 4, 10, 7	0.138	0.065	2.118	

\*These clusters are internal to the less preferred cluster #4.

\*\*Includes a nearby but not overlapping epitope.

**Table 46**  
**BIMAS-NIH/Parker algorithm Results for PSCA**

Rank	Start	Score
1	43	153
2	5	84
3	7	79
4	109	36
5	105	105
6	108	24
7	14	21
8	20	18
9	115	17
10	42	15
11	36	15
12	99	9
13	58	8
		20

25 **Table 47**  
**SYFPEITHI (Rammensee algorithm) Results for PSCA**

Rank	Start	Score	Rank	Start	Score
1	108	30	17	54	19
2	14	30	18	12	19
3	105	29	19	4	19
4	5	28	20	1	19
5	115	26	21	112	18
6	99	26	22	101	18
7	7	26	23	98	18
8	109	24	24	51	18
9	53	23	25	43	18
10	107	21	26	106	17
11	20	21	27	104	17
12	8	21	28	83	17
13	13	20	29	63	17
14	102	19	30	50	17
15	60	19	31	3	17
16	57	19	32	9	16
			33	92	16

Table 48

## Prediction of clusters for prostate stem cell antigen (PSCA)

Total AAs: 123

Total 9-mers: 115

SYFPEITHI 16: 33;

SYFPEITHI 20: 13

NIH 5: 13

			Epitopes/AA			
	Cluster #	AAs	Epitopes (by rank)	Cluster	Whole Pr.	Ratio
SYFPEITHI >16	1	1 to 28	20, 31, 19, 4, 7, 12, 33, 18, 13, 2, 11	0.393	0.268	1.464
	2	43-71	25, 30, 24, 9, 17, 16, 15, 29	0.276	0.268	1.028
	3	92-123	32, 23, 6, 27, 14, 22, 3, 26, 10, 1, 8, 21, 5	0.406	0.268	1.514
SYFPEITHI >20	1	5 to 28	4, 7, 12, 13, 2, 11	0.250	0.106	2.365
	2	99-123	6, 3, 10, 1, 8, 5	0.240	0.106	2.271
NIH	1	5 to 28	2, 3, 7, 8	0.167	0.106	1.577
	2	36-51	11, 10, 1	0.188	0.106	1.774
	3	99-123	12, 5, 6, 4, 9	0.200	0.106	1.892
	4*	105-116	5, 6, 4	0.250	0.106	2.365

\*This cluster is internal to the less preferred cluster #3.

In tables 49-60 epitope prediction and cluster analysis data for each algorithm are presented together in a single table.

**Table 49****Prediction of clusters for MAGE-1 (NIH algorithm)**

Total AAs: 309

Total 9-mers: 301

NIH 5:19 9-mers

Cluster #	AAs	Epitope Rank	Start Position	NIH Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	18-32	16	18	9	0.133	0.063	2.112
		19	24	7			
2	101-113	14	101	11	0.154	0.063	2.442
		7	105	44			
3	146-159	9	146	32	0.143	0.063	2.263
		3	151	169			
4	169-202	10	169	32	0.176	0.063	2.796
		13	174	16			
		18	181	8			
		17	187	8			
		6	188	74			
		5	194	110			
5	264-277	2	264	190	0.143	0.063	2.263
		12	269	20			
6	278-290	1	278	743	0.154	0.063	2.437
		11	282	28			

**Table 50****Prediction of clusters for MAGE-1 (SYFPEITHI algorithm)**

Total AAs: 309

Total 9-mers: 301

SYFPEITHI 16: 46 9-mers

Cluster #	Aas	Epitope Rank	Start Position	SYFPEITHI Score	Epitopes/AA		
					Cluster	Whole	Ratio
1	7-49	22	7	19	0.233	0.153	1.522
		9	15	22			
		27	18	18			
		16	20	20			
		28	22	18			
		29	24	18			
		33	31	17			
		30	35	18			
		2	38	26			
		17	41	20			
2	89-132	10	89	22	0.273	0.153	1.783
		18	92	20			
		7	93	23			
		23	96	19			
		43	98	16			
		4	101	25			
		8	105	23			
		34	107	17			
		35	108	17			
		36	113	17			
		37	118	17			
		19	124	20			
3	167-203	44	167	16	0.270	0.153	1.766
		20	169	20			
		12	174	21			
		24	181	19			
		6	187	24			
		31	188	18			
		25	191	19			
		38	192	17			
		1	194	27			
		13	195	21			
		4	230	21			
		39	238	17			
5	264-297	15	264	21	0.235	0.153	1.538
		32	269	18			
		40	270	17			
		26	271	19			
		46	275	16			
		3	278	26			
		21	282	20			
		41	289	17			

**Table 51**  
**Prediction of clusters for MAGE-2 (NIH algorithm)**

Total AAs: 314  
 Total 9-mers: 308  
 NIH >= 5: 20 9-mers

Cluster #	AAs	Epitope Rank	Start Position	NIH Score	Cluster	Epitope/AA Whole Pr.	Ratio
1	101-120	18	101	5.373	0.150	0.065	2.310
		16	108	6.756			
		1	112	2800.697			
2	153-167	8	153	31.883	0.200	0.065	3.080
		4	158	168.552			
		7	159	32.138			
3	169-211	14	169	8.535	0.209	0.065	3.223
		19	174	5.346			
		6	176	49.993			
		11	181	15.701			
		15	188	7.536			
		12	195	12.809			
		5	200	88.783			
		10	201	16.725			
		17	203	5.609			
4	271-284	3	271	398.324	0.143	0.065	2.200
		9	276	19.658			

Table 52

## Prediction of clusters for MAGE-2 (SYFPEITHI algorithm)

Total AAs: 314

Total 9-mers: 308

SYFPEITHI 16: 52 9-mers

Cluster #	AAs	Epitope Rank	Start Position	SYFPEITHI Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	15-32	13	15	21	0.278	0.169	1.645
		29	18	18			
		43	20	16			
		30	22	18			
		21	24	19			
2	37-56	31	37	18	0.250	0.169	1.481
		16	40	20			
		44	44	16			
		14	45	21			
		22	48	19			
3	96-133	36	96	17	0.211	0.169	1.247
		46	101	16			
		6	108	25			
		47	109	16			
		2	112	27			
		37	120	17			
		38	125	17			
		17	131	20			
4	153-216	12	153	22	0.344	0.169	2.036
		39	158	17			
		7	159	25			
		23	161	19			
		24	162	19			
		48	164	16			
		49	167	16			
		32	170	18			
		50	171	16			
		4	174	26			
		9	176	24			
		51	177	16			
		15	181	21			
		25	188	19			
		18	194	20			
		33	195	18			
		19	198	20			
		3	200	27			
		1	201	28			
		40	202	17			
		10	203	23			
		52	208	16			
5	237-254	26	237	19	0.167	0.169	0.987
		27	245	19			
		34	246	18			
6	271-299	8	271	25	0.241	0.169	1.430
		35	276	18			
		41	277	17			
		11	278	23			
		28	283	19			
		20	285	20			
		42	291	17			

**Table 53****Prediction of clusters for MAGE-3 (NIH algorithm)**

Total AAs: 314

Total 9-mers: 308

NIH 5: 22 9-mers

Cluster #	AAs	Epitope Rank	Start Position	NIH Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	101-120	15	101	11.002	0.200	0.071	2.800
		21	105	6.488			
		8	108	49.134			
		2	112	339.313			
2	153-167	18	153	7.776	0.200	0.071	2.800
		6	158	51.77			
		22	159	5.599			
3	174-209	17	174	8.832	0.194	0.071	2.722
		7	176	49.993			
		13	181	15.701			
		19	188	7.536			
		14	195	12.809			
		5	200	88.783			
		12	201	16.725			
4	237-251	16	237	10.868	0.200	0.071	2.800
		4	238	148.896			
		20	243	6.88			
5	271-284	1	271	2655.495	0.143	0.071	2.000
		11	276	19.658			

Table 54

## Prediction of clusters for MAGE-3 (SYFPEITHI algorithm)

Total AAs: 314

Total 9-mers: 308

SYFPEITHI 16: 47 9-mers

Cluster #	AAs	Epitope Rank	Start Position	SYFPEITHI Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	15-32	12	15	21	0.278	0.153	1.820
		26	18	18			
		37	20	16			
		27	22	18			
		18	24	19			
2	38-56	38	38	16	0.263	0.153	1.725
		15	40	20			
		39	44	16			
		13	45	21			
		19	48	19			
3	101-142	28	101	18	0.190	0.153	1.248
		40	105	16			
		1	108	31			
		6	112	25			
		31	120	17			
		32	125	17			
		16	131	20			
		41	134	16			
4	153-216	20	153	19	0.313	0.153	2.048
		29	156	18			
		33	158	17			
		21	159	19			
		34	161	17			
		42	164	16			
		43	167	16			
		10	174	22			
		8	176	23			
		14	181	21			
		22	188	19			
		44	193	16			
		11	194	22			
		23	195	19			
		45	197	16			
		17	198	20			
		3	200	27			
		2	201	28			
		35	202	17			
		46	208	16			
5	220-230	5	220	26	0.182	0.153	1.191
		47	222	16			
6	237-246	7	237	25	0.200	0.153	1.311
		9	238	23			
7	271-293	4	271	27	0.217	0.153	1.425
		30	276	18			
		24	278	19			
		36	283	17			
		25	285	19			

**Table 55**  
**Prediction of clusters for PRAME (NIH algorithm)**

Total AAs: 509  
 Total 9-mers: 501  
 NIH 5: 40 9-mers

Cluster #	AAs	Epitope Rank	Start Position	NIH Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	33-47	20	33	18	0.133	0.080	1.670
		17	39	21			
2	71-81	9	71	50	0.2	0.07984	2.505
		32	73	7			
3	99-108	23	100	15	0.2	0.07984	2.505
		24	99	13			
4	126-135	38	126	5	0.2	0.07984	2.505
		35	127	6			
5	224-246	5	224	124	0.130	0.080	1.634
		8	230	63			
		39	238	5			
6	290-303	18	290	18	0.214	0.080	2.684
		14	292	23			
		7	295	66			
7	305-324	28	305	10	0.200	0.080	2.505
		30	308	8			
		25	312	13			
		36	316	6			
8	394-409	2	394	182	0.188	0.080	2.348
		12	397	42			
		31	401	7			
9	422-443	10	422	49	0.227	0.080	2.847
		3	425	182			
		34	431	7			
		29	432	9			
		4	435	160			
10	459-487	15	459	21	0.172	0.080	2.159
		11	462	45			
		22	466	15			
		40	472	5			
		37	479	6			

**Table 56**  
**Prediction of clusters for PRAME (SYFPEITHI algorithm)**

Total AAs: 509

Total 9-mers: 501

SYFPEITHI 17: 80 9-mers

Cluster #	AAs	Epitope Rank	Start Position	SYFPEITHI Score	Cluster	Epitopes/AA Whole Pr.	Ratio
1	18-59	65	18	17	0.238	0.160	1.491
		50	21	18			
		66	26	17			
		35	33	20			
		22	34	22			
		51	37	18			
		5	39	27			
		23	40	22			
		13	44	24			
		46	51	19			
2	78-115	36	78	20	0.263	0.160	1.648
		67	80	17			
		52	84	18			
		24	86	22			
		53	91	18			
		25	93	22			
		9	99	25			
		8	100	26			
		54	103	18			
		55	107	18			
3	191-202	56	191	18	0.167	0.160	1.044
		38	194	20			
4	205-215	26	205	22	0.182	0.160	1.139
		27	207	22			
5	222-238	47	222	19	0.235	0.160	1.474
		14	224	24			
		69	227	17			
		57	230	18			
6	241-273	70	241	17	0.212	0.160	1.328
		15	248	24			
		71	255	17			
		30	258	21			
		39	259	20			
		58	261	18			
		40	265	20			
7	290-342	72	290	17	0.208	0.160	1.300
		48	293	19			
		31	298	21			
		73	301	17			
		18	305	23			
		6	308	27			
		10	312	25			
		19	316	23			
		28	319	22			

**Prediction of clusters for PRAME (SYFPEITHI algorithm)**

Total AAs: 509

Total 9-mers: 501

SYFPEITHI 17: 80 9-mers

Cluster #	AAs	Epitope Rank	Start Position	SYFPEITHI Score	Cluster	Epitopes/AA Whole Pr.	Ratio
8	343-363	41	326	20	0.238	0.160	1.491
		74	334	17			
		59	343	18			
		60	348	18			
		75	351	17			
		20	353	23			
9	364-447	76	355	17	0.250	0.160	1.566
		49	364	19			
		32	371	21			
		11	372	25			
		61	375	18			
		77	382	17			
		21	390	23			
		78	391	17			
		1	394	30			
		42	397	20			
		62	403	18			
		33	410	21			
		43	418	20			
		34	419	21			
		7	422	27			
		2	425	29			
		79	426	17			
		63	428	18			
		64	431	18			
		12	432	25			
		16	435	24			
		80	439	17			
10	455-474	29	455	22	0.200	0.160	1.253
		17	459	24			
		4	462	28			
		3	466	29			

**Table 57**  
**Predication of clusters for CEA (NIH algorithm)**

Total AAs:702  
 Total 9-mers: 694  
 NIH 5: 30 9-mers

Cluster #	AA	Peptides	Start	Score	Cluster	Peptides/AAs	Whole Pr.	Ratio
		Rank	Position					
1	17-32	5	17	79.041	0.188	0.043	4.388	
		7	18	46.873				
		20	24	12.668				
2	113-129	2	113	167.991	0.118	0.043	2.753	
		15	121	21.362				
3	172-187	25	172	9.165	0.125	0.043	2.925	
		14	179	27.995				
4	278-291	30	278	5.818	0.143	0.043	3.343	
		17	283	19.301				
5	350-365	9	350	43.075	0.125	0.043	2.925	
		12	357	27.995				
6	528-543	8	528	43.075	0.125	0.043	2.925	
		13	535	27.995				
7	631-645	23	631	9.563	0.200	0.043	4.680	
		19	634	13.381				
		24	637	9.245				
8	691-702	1	691	196.407	0.167	0.043	3.900	
		27	694	7.769				

**Table 58**  
**Predication of clusters for CEA (SYFPEITHI algorithm)**

Total AAs:702

Total 9-mers: 694

SYFPEITHI 16: 81 9-mers

Cluster #	AA	Peptides Rank	Start Position	Score	Cluster	Peptides/AAs Whole Pr.	Ratio
1	5-36	67	5	16	0.250	0.117	2.140
		23	12	19			
		24	16	19			
		9	17	22			
		25	18	19			
		32	19	18			
		68	23	16			
		33	28	18			
2	37-62	41	37	17	0.269	0.117	2.305
		20	44	20			
		26	45	19			
		42	46	17			
		27	50	19			
		43	53	17			
		44	54	17			
3	99-115	14	99	21	0.235	0.117	2.014
		5	100	23			
		45	104	17			
		34	107	18			
4	116-129	69	116	16	0.143	0.117	1.223
		21	121	20			
5	172-187	46	172	17	0.125	0.117	1.070
		70	179	16			
6	192-202	3	192	24	0.182	0.117	1.557
		47	194	17			
7	226-241	48	226	17	0.188	0.117	1.605
		49	229	17			
		15	233	21			
8	307-318	11	307	22	0.250	0.117	2.140
		71	308	16			
		51	310	17			
9	319-349	52	319	17	0.129	0.117	1.105
		53	327	17			
		72	335	16			
		35	341	18			
10	370-388	12	370	22	0.211	0.117	1.802
		54	372	17			
		74	375	16			
		6	380	23			
11	403-419	56	403	17	0.235	0.117	2.014
		57	404	17			
		58	407	17			
		28	411	19			

12	427-442	59	427	17	0.188	0.117	1.605
		75	432	16			
		76	434	16			
13	450-462	77	450	16	0.154	0.117	1.317
		13	454	22			
14	488-505	36	488	18	0.167	0.117	1.427
		18	492	21			
		60	497	17			
15	548-558	4	548	24	0.182	0.117	1.557
		61	550	17			
16	565-577	62	565	17	0.154	0.117	1.317
		19	569	21			
17	579-597	78	579	16	0.143	0.117	1.223
		79	582	16			
		7	589	23			
18	605-618	2	605	25	0.143	0.117	1.223
		38	610	18			
19	631-669	29	631	19	0.154	0.117	1.317
		63	637	17			
		80	644	16			
		64	652	17			
		39	660	18			
		81	661	16			
20	675-702	22	675	20	0.286	0.117	2.446
		30	683	19			
		31	687	19			
		40	688	18			
		65	690	17			
		1	691	31			
		66	692	17			
		8	694	23			

**Table 59**  
**Predication of clusters for SCP-1 (NIH algorithm)**

Total AAs: 976  
 Total 9-mers: 968  
 NIH 5: 37 9-mers

Cluster #	AA	Peptides Rank	Start Position	Score	Cluster	Peptides/AAs Whole Pr.	Ratio
1	101-116	15	101	40.589	0.125	0.038	3.270
		13	108	57.255			
2*	281-305	14	281	44.944	0.12	0.038	3.139
		24	288	15.203			
		17	297	32.857			
3	431-447	8	431	80.217	0.073	0.038	1.914
		26	438	11.861			
		4	439	148.896			
4	557-579	11	557	64.335	0.174	0.038	4.550
		19	560	24.937			
		6	564	87.586			
		18	571	32.765			
5	635-650	10	635	69.552	0.125	0.038	3.270
		34	642	6.542			
6	755-767	36	755	5.599	0.154	0.038	4.025
		35	759	5.928			
7	838-854	2	838	284.517	0.118	0.038	3.078
		28	846	11.426			

**Table 60**  
**Predication of clusters for SCP-1**

Total AAs: 976

Total 9-mers: 968

Rammensee 16: 118 9-mers

Cluster #	AA	Peptides Rank	Start Position	Score	Cluster	Peptides/AAs Whole Pr.	Ratio
1	8-28	99	8	16	0.143	0.121	1.182
		77	15	17			
		100	20	16			
2	63-80	78	63	17	0.222	0.121	1.838
		50	66	19			
		102	69	16			
		60	72	18			
3	94-123	79	94	17	0.133	0.121	1.103
		12	101	23			
		17	108	22			
		103	115	16			
4	126-158	35	126	20	0.182	0.121	1.504
		36	133	20			
		51	139	19			
		80	140	17			
		61	143	18			
		37	150	20			
5	161-189	38	161	20	0.207	0.121	1.711
		52	165	19			
		81	171	17			
		82	177	17			
		62	178	18			
		39	181	20			
6	213-230	40	213	20	0.167	0.121	1.379
		13	220	23			
		28	222	21			
7	235-250	63	235	18	0.125	0.121	1.034
		18	242	22			
8	260-296	83	260	17	0.243	0.121	2.012
		105	262	16			
		84	267	17			
		106	269	16			
		41	270	20			
		64	271	18			
		85	274	17			
		19	281	22			
		3	288	25			
9	312-338	108	312	16	0.148	0.121	1.225
		29	319	21			
		30	323	21			
		65	330	18			
10	339-355	66	339	18	0.235	0.121	1.946
		31	340	21			
		42	344	20			
		53	347	19			

11	376-447	54	376	19	0.194	0.121	1.608
		43	382	20			
		44	386	20			
		20	390	22			
		55	397	19			
		6	404	24			
		86	407	17			
		45	411	20			
		67	417	18			
		21	425	22			
		46	431	20			
		68	432	18			
		32	438	21			
		7	439	24			
12	455-488	33	455	21	0.235	0.121	1.946
		47	459	20			
		56	462	19			
		87	463	17			
		88	466	17			
		14	470	23			
		109	473	16			
		34	480	21			
13	515-530	57	515	19	0.125	0.121	1.034
		22	522	22			
14	557-590	8	557	24	0.147	0.121	1.216
		23	564	22			
		9	571	24			
		90	575	17			
		58	582	19			
15	610-625	69	610	18	0.125	0.121	1.034
		91	617	17			
16	633-668	92	633	17	0.222		
		10	635	24			
		70	638	18			
		93	640	17			
		48	642	20			
		49	645	20			
		111	652	16			
		112	660	16			
17	674-685	71	674	18	0.167	0.121	1.379
		11	677	24			
18	687-702	1	687	26	0.125	0.121	1.034
		94	694	17			
19	744-767	113	744	16	0.250	0.121	2.068
		95	745	17			
		4	745	25			
		24	752	22			
		2	755	26			
		72	759	18			
20	812-827	97	812	17	0.125	0.121	1.034
		115	819	16			
21	838-857	116	838	16	0.150	0.121	1.241
		25	846	22			
		74	849	18			

22	896-913	117	896	16	0.222	0.121	1.838
		98	899	17			
		26	902	22			
		76	905	18			

1 MLLAVLYCLL WSFQTSAGHF PRACVSSKNL MEKECCPPWS GDRSPCGQLS  
 GRGSCQNILL  
 5 61 SNAPLGPQFP FTGVDDRESW PSVFYNRTCQ CSGNFMGFNC GNCKFGFWGP  
 NCTERRLLVR  
 121 RNIFDLSAPE KDKFFAYLTL AKHTISSLDVY IPIGTYGQMK NGSTPMFNDI  
 NIYDLFVWMH  
 10 181 YYVSMDALLG GSEIWRDIDF AHEAPAFLPW HRLFLLRWEQ EIQLTGDEN  
 FTIPYWDWRD  
 241 AEKCDICTDE YMGGQHPTNP NLLSPASFFS SWQIVCSRLE EYNSHQSLCN  
 GTPEGPLRRN  
 301 PGNHDKSRTP RLPSSADVEF CLSLTQYESG SMDKAANFSF RNTLEGFASP  
 LTGIADASQS  
 15 361 SMHNALHIYM NGTMSQVQGS ANDPIFLHH AFVDSIFEQW LRRHRPLQEV  
 YPEANAPIGH  
 421 NRESYMPFFI PLYRNGDFFI SSKDLGYDYS YLQDSDPDSF QDYIKSYLEQ  
 ASRIWSWLLG  
 20 481 AAMVGAVLTA LLAGLVSLLC RHKRKQLPEE KQPLLMEKED YHSLYQSHL  
 TYROSINASE PROTEIN

25

30 1 MNGDDAFARR PTVGAQIPEK IQKAFDDIAK YFSKEEWEKM KASEKIFYVY  
 MKRKYEAMTK  
 61 LGFKATLPPF MCNKRAEDFQ GNDLDNDPNR GNQVERPQMT FGRLQGISPK  
 IMPKKPAEEG  
 35 121 NDSEEVPEAS GPQNDGKELC PPGKPTTSEK IHERSGPKRG EHAWTHRLRE  
 RKQLVIYEEI  
 181 SDPEEDDE  
 SSX-2 PROTEIN

40

1 MWNLLHETDS AVATARRPRW LCAGALVLAG GFFLLGFLFG WFIKSSNEAT  
 NITPKHNMKA  
 5 61 FLDELKAENI KKFLYNFTQI PHLAGTEQNF QLAKQIQSQW KEFGLDSVEL  
 AHYDVLLSYP  
 121 NKTHPNYISI INEDGNEIFN TSLFEPFFFFP YENVSDIVPP FSAFSPQGMP  
 EGDLIVVNYA  
 181 RTEDFFKLER DMKINCSGKI VIARYGKVFR GNKVKNAQLA GAKGVILYSD  
 PADYFAPGVK  
 10 241 SYPDGVNLPG GGVQRGNILN LNGAGDPLTP GYPANEYAYR RGIAEAVGLP  
 SIPVHPIGYY  
 301 DAQKLLEKMG GSAPPDSSWR GSLKVPYNVG PGFTGNFSTQ KVKMHIHSTN  
 EVTRIYNVIG  
 361 TLRGAVEPDR YVILGGHRDS WVFGGIDPQS GAAVVHEIVR SFGTLKKEGW  
 15 RPRRTILFAS  
 421 WDAEEFGLLG STEWAEENSR LLQERGVAYI NADSSIEGNY TLRVDCTPLM  
 YSLVHNLTKE  
 481 LKSPDEGFEG KSLYESWTKK SPSPEFSGMP RISKLGSGND FEVFFQRLGI  
 ASGRARYTKN  
 20 541 WETNKFSGYP LYHSVYETYE LVEKFYDPMF KYHLTVAQVR GGMVFELANS  
 IVLPFDCRDY  
 601 AVVLRKYADK IYSISMKHPQ EMKTYSVSFD SLFSAVKNFT EIASKFSERL  
 QDFDKSNPIV  
 661 LRMMNDQLMF LERAFIDPLG LPDRPFYRHV IYAPSSHNKY AGESFPGIYD  
 25 ALFDIESKVD  
 721 PSKAWGEVKR QIYVAAFTVQ AAAETLSEVA

## PSMA PROTEIN

30

1:	Homo sapiens tyrosinase	<u>PubMed</u> , <u>Protein</u> ,
	NM_00037 (oculocutaneous albinism IA) (TYR),	<u>Taxonomy</u> , <u>OMIM</u> , <u>LinkOut</u>
2	Mrna	
LOCUS	NM_000372 1964 bp mRNA	PRI 31-
OCT-2000		
DEFINITION	Homo sapiens tyrosinase (oculocutaneous albinism IA)	
(TYR), mRNA.		
ACCESSION	NM_000372	
VERSION	NM_000372.1 GI:4507752	
KEYWORDS	.	
SOURCE	human.	
40 ORGANISM	<u>Homo sapiens</u>	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;	
	Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae;	
	Homo.	
45 REFERENCE	1 (bases 1 to 1964)	
	AUTHORS Kwon BS, Haq AK, Pomerantz SH and Halaban R.	
	TITLE Isolation and sequence of a cDNA clone for human	
	tyrosinase that	
	maps at the mouse c-albino locus	
50 JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 84 (21), 7473-7477	
(1987)		
MEDLINE	88041128	
PUBMED	2823263	
REMARK	Erratum: [[published erratum appears in Proc Natl Acad	
55 Sci U S A		

1988 Sep;85(17):6352]]  
 REFERENCE 2 (bases 1 to 1964)  
 AUTHORS Barton DE, Kwon BS and Francke U.  
 TITLE Human tyrosinase gene, mapped to chromosome 11 (q14---  
 5 -q21), defines second region of homology with mouse  
 chromosome 7  
 JOURNAL Genomics 3 (1), 17-24 (1988)  
 MEDLINE 89122007  
 PUBMED 3146546  
 REFERENCE 3 (bases 181 to 1964)  
 AUTHORS Shibahara,S., Tomita,Y., Tagami,H., Muller,R.M. and  
 Cohen,T.  
 TITLE Molecular basis for the heterogeneity of human  
 15 tyrosinase  
 JOURNAL Tohoku J. Exp. Med. 156 (4), 403-414 (1988)  
 MEDLINE 89222868  
 REFERENCE 4 (bases 1 to 1964)  
 AUTHORS Bouchard B, Fuller BB, Vijayasaradhi S and Houghton  
 20 AN.  
 TITLE Induction of pigmentation in mouse fibroblasts by  
 expression of  
 human tyrosinase cDNA  
 JOURNAL J. Exp. Med. 169 (6), 2029-2042 (1989)  
 MEDLINE 89279151  
 PUBMED 2499655  
 REFERENCE 5 (bases 1 to 1964)  
 AUTHORS Takeda,A., Tomita,Y., Okinaga,S., Tagami,H. and  
 Shibahara,S.  
 30 TITLE Functional analysis of the cDNA encoding human  
 tyrosinase precursor  
 JOURNAL Biochem. Biophys. Res. Commun. 162 (3), 984-990 (1989)  
 MEDLINE 89351001  
 REFERENCE 6 (bases 1 to 1964)  
 AUTHORS Kikuchi H, Miura H, Yamamoto H, Takeuchi T, Dei T and  
 35 Watanabe M.  
 TITLE Characteristic sequences in the upstream region of the  
 human tyrosinase gene  
 40 JOURNAL Biochim. Biophys. Acta 1009 (3), 283-286 (1989)  
 MEDLINE 90089403  
 PUBMED 2480811  
 REFERENCE 7 (bases 1 to 1964)  
 AUTHORS Giebel LB, Strunk KM and Spritz RA.  
 45 TITLE Organization and nucleotide sequences of the human  
 tyrosinase gene  
 and a truncated tyrosinase-related segment  
 JOURNAL Genomics 9 (3), 435-445 (1991)  
 MEDLINE 91236163  
 PUBMED 1903356  
 REFERENCE 8 (bases 1 to 1964)  
 AUTHORS Brichard V, Van Pel A, Wolfel T, Wolfel C, De Plaen E,  
 Lethe B,  
 Coulie P and Boon T.  
 55 TITLE The tyrosinase gene codes for an antigen recognized by  
 autologous cytolytic T lymphocytes on HLA-A2 melanomas  
 JOURNAL J. Exp. Med. 178 (2), 489-495 (1993)

MEDLINE 93340625  
 PUBMED 8340755  
 COMMENT PROVISIONAL REFSEQ: This record has not yet been  
 subject to final  
 5 NCBI review. The reference sequence was derived from  
 M27160.1.  
 FEATURES Location/Qualifiers  
 source 1..1964  
 10 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11q14-q21"  
 /cell\_line="MeWo melanoma DNA, and cDNA to  
 15 mRNA"  
 /tissue\_type="placenta"  
 gene 1..1964  
 /gene="TYR"  
 /note="OCAIA"  
 /db\_xref="LocusID:7299"  
 20 /db\_xref="MIM:203100"  
 CDS 83..1672  
 /gene="TYR"  
 /EC\_number="1.14.18.1"  
 /note="Tyrosinase"  
 25 /codon\_start=1  
 /db\_xref="LocusID:7299"  
 /db\_xref="MIM:203100"  
 /product="tyrosinase (oculocutaneous albinism  
 30 IA)"  
 /protein\_id="NP\_000363.1"  
 /db\_xref="GI:4507753"  
 /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLMEKECCPPWSGDRS  
 35 PCGQLSGRGSCQNIILSNAPLGPQFPFTGVDDRESWPSVFYNRTQCSGNFMGFCGN  
 CKFGFWGPNCTERLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSLQYVPIGTYGQMK  
 NGSTPMFNDINIYDLFVWMHYYSMDALLGGSEIWRDIDFAHEAPAFLPWHRLFLLRW  
 40 EQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGQHPTNPNLSPASFFSSWQIVC  
 SRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPLPSSADVEFCLSLTQYESGSMDK  
 45 AANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQVQGSANDPIFLHHAF  
 VDSIFEQWLRRRPLQEVYPEANAPIGHNRESYMPFIPLYRNGDFFISSKDLGYDYS  
 YLQDSDPDSFQDYIKSYLEQASRIWSLLGAAMVGAVLTALLAGLVSLLCRHKRKQLP  
 50 EEKQPLLMKEKDYHSLYQSHL"  
 sig peptide 83..136  
 mat peptide 137..1669  
 /note="tyrosinase"  
 /product="tyrosinase (oculocutaneous albinism  
 55 IA)"  
 misc feature 560..1429  
 /note="tyrosinase; Region: Common central  
 domain of

"tyrosinase"

BASE COUNT	520	a	462	c	445	g	537	t
ORIGIN								

5            1 atcactgtag tagtagctgg aaagagaaaat ctgtgactcc aattagccag  
 ttcctgcaga  
       61 ccttgcagg actagaggaa gaatgctct ggctgttttgc tactgcctgc  
 tgtggagttt  
       121 ccagacccctcc gctggccatt tccctagagc ctgtgtctcc tctaagaacc  
 tcatggagaa  
 10        181 ggaatgctgt ccaccgtggaa gggggacag gagtcctgtt ggccagcttt  
 caggcagagg  
       241 ttccctgtcag aatatcccttc tgtccaatgc accacttggg cctcaatttc  
 cttcacagg  
       301 ggtggatgac cgggagtcgt ggcttccgtt ctttataat aggacctgcc  
 agtgctctgg  
 15        361 caacttcatg ggattcaact gtggaaactg caagtttggc ttttggggac  
 caaactgcac  
       421 agagagacga ctcttggtga gaagaaaacat cttcgatttgc agtgcggcag  
 agaaggacaa  
 20        481 atttttgcc tacctcactt tagcaaagca taccatcagc tcagactatg  
 tcataccccat  
       541 agggacctat ggccaaatga aaaatggatc aacacccatg tttaacgaca  
 tcaatattta  
       601 tgacctcttt gtctggatgc attattatgt gtcaatggat gcactgcttg  
 ggggatctga  
       661 aatctggaga gacattgatt ttgccccatga agcaccagct tttctgcctt  
 ggcatagact  
       721 ctcttggttt cgggtggaaac aagaaatcca gaagctgaca ggagatgaaa  
 acttcactat  
 30        781 tccatattgg gactggcggg atgcagaaaa gtgtgacatt tgcacagatg  
 agtacatggg  
       841 aggtcagcac cccacaaatc ctaacttact cagcccagca tcattttct  
 cctcttggca  
       901 gattgtctgt agccgattgg aggagtacaa cagccatcag tctttatgca  
 atggAACGCC  
       961 cgagggacct ttacggcgta atcctggaaa ccatgacaaaa tccagaaccc  
 caaggtcccc  
       1021 ctcttcagct gatgtagaat tttgcctgag tttgacccaa tatgaatctg  
 gttccatgg  
 40        1081 taaaagctgcc aatttcagct ttagaaatac actggaaggaa tttgctagtc  
 cacttactgg  
       1141 gatagcggat gcctctcaaa gcagcatgca caatgccttg cacatctata  
 tgaatggac  
 45        1201 aatgtcccag gtacaggat ctgccaacga tcctatcttc cttcttcacc  
 atgcatttgt  
       1261 tgacagtatt tttgagcagt ggctccgaag gcaccgtcct cttcaagaag  
 tttatccaga  
       1321 agccaatgca cccattggac ataaccggga atcctacatg gttcctttta  
 taccactgta  
 50        1381 cagaaatggt gatttcttta tttcatccaa agatctgggc tatgactata  
 gctatctaca  
       1441 agattcagac ccagactttt ttcaagacta cattaagtcc tattttggaaac  
 aagcgagtcg  
       1501 gatctggtca tggctcccttggggcgat ggttaggggcc gtcctcactg  
 ccctgctggc  
       1561 agggcttggtgc agcttgctgt gtctgcacaa gagaaagcag cttcctgaag  
 aaaagcagcc

1621 actcctcatg gagaaagagg attaccacag cttgtatcag agccattat  
 aaaaggctta  
 1681 ggcaatagag tagggccaaa aagctgacc tcactctaac tcaaagtaat  
 5 gtccagggtc  
 1741 ccagagaata tctgctggta ttttctgta aagaccattt gcaaaattgt  
 aacctaatac  
 1801 aaagtgtac cttttccaa ctcaggtaga acacacctgt ctttgttgc  
 ctgtttcac  
 1861 tcagccctt taacatttc ccctaagccc atatgtctaa ggaaaggatg  
 10 ctatggta  
 1921 atgaggaact gttatttgc tgtgaattaa agtgcttta tttt

		<u>PubMed, Protein, Related Sequences, Taxonomy, OMIM, LinkOut</u>
	1: Homo sapiens synovial sarcoma, NM_0031 X breakpoint 2 (SSX2), mRNA	47
15	LOCUS NM_003147 766 bp mRNA	PRI 14-
	MAR-2001	
	DEFINITION Homo sapiens synovial sarcoma, X breakpoint 2 (SSX2), mRNA.	
20	ACCESSION NM_003147	
	VERSION NM_003147.1 GI:10337582	
	KEYWORDS	
	SOURCE human.	
	ORGANISM Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;	
25	Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae;	
	Homo.	
	REFERENCE 1 (bases 1 to 766)	
	AUTHORS Shipley JM, Clark J, Crew AJ, Birdsall S, Rocques PJ,	
30	Gill S,	
	Chelly J, Monaco AP, Abe S, Gusterson BA and et al.	
	TITLE The t(X;18)(p11.2;q11.2) translocation found in human	
	synovial	
	sarcomas involves two distinct loci on the X	
35	chromosome	
	JOURNAL Oncogene 9 (5), 1447-1453 (1994)	
	MEDLINE 94203675	
	PUBMED 8152806	
	REFERENCE 2 (bases 1 to 766)	
40	AUTHORS Crew,A.J., Clark,J., Fisher,C., Gill,S., Grimer,R.,	
	Chand,A.,	
	Shipley,J., Gusterson,B.A. and Cooper,C.S.	
	TITLE Fusion of SYT to two genes, SSX1 and SSX2, encoding	
	proteins with	
45	synovial sarcoma	
	JOURNAL EMBO J. 14 (10), 2333-2340 (1995)	
	MEDLINE 95292974	
	REFERENCE 3 (bases 1 to 766)	
50	AUTHORS Tureci O, Sahin U, Schobert I, Koslowski M, Scmitt H, Schild HJ,	
	Stenner F, Seitz G, Rammensee HG and Pfreundschuh M.	
	TITLE The SSX-2 gene, which is involved in the t(X;18)	
	translocation of	

synovial sarcomas, codes for the human tumor antigen

HOM-MEL-40  
 JOURNAL      Cancer Res. 56 (20), 4766-4772 (1996)  
 MEDLINE      96438636  
 PUBMED      8840996  
 5      COMMENT      PROVISIONAL REFSEQ: This record has not yet been  
 subject to final  
 NCBI review. The reference sequence was derived from  
 X86175.1.  
 10      FEATURES      Location/Qualifiers  
 source      1..766  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 15      gene      1..766  
 /map="Xp11.23-p11.22"  
 /dev\_stage="adult"  
 /gene="SSX2"  
 /note="HD21; HOM-MEL-40; SSX"  
 /db\_xref="LocusID:6757"  
 /db\_xref="MIM:300192"  
 20      misc feature      20..61  
 /note="Kruppel associated box homology"  
 CDS      92..658  
 /gene="SSX2"  
 /codon\_start=1  
 /db\_xref="LocusID:6757"  
 /db\_xref="MIM:300192"  
 /product="synovial sarcoma, X breakpoint 2"  
 /protein\_id="NP\_003138.1"  
 30      /db\_xref="GI:10337583"  
  
 /translation="MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEEWEKMKASE  
 35      KIFYVYMKRKYEAMTKLGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFG  
 RLQGISPKIMPKKPAEEGNDSEEVPEASGPQNDGKELCPPGKPTTSEKIHERSGPKRG  
 EHAWTHRLRERKQLVIYEEISDPEEDDE"  
 misc feature      161..337  
 40      /note="KRAB; Region: krueppel associated box"  
 BASE COUNT      229 a      181 c      200 g      156 t  
 ORIGIN  
 1 ctctctttcg attcttccat actcagagta cgcacggct gattttctct  
 ttggattctt  
 45      61 cccaaatcag agtcagactg ctcccggtgc catgaacgga gacgacgcct  
 ttgcaaggag  
 121 acccacggtt ggtgctaaa taccagagaa gatccaaaag gccttcgatg  
 atattgccaa  
 181 atacttctct aaggaagagt gggaaaagat gaaagcctcg gagaaaatct  
 50      tctatgtata  
 241 tatgaagaga aagtatgagg ctatgactaa actaggtttc aaggccaccc  
 tcccacccctt  
 301 catgtgtaat aaacggggccg aagacttcca gggaaatgat ttggataatg  
 accctaaccg  
 55      361 tggaaatcag gttgaacgtc ctcagatgac tttcggcagg ctccaggaa  
 tctcccccga  
 421 gatcatgccc aagaagccag cagaggaagg aaatgattcg gaggaagtgc  
 cagaagcatac

481 tggcccacaa aatgatggga aagagctgtg ccccccggga aaaccaacta  
 cctctgagaa  
 5 541 gattcacgag agatctggac ccaaaagggg ggaacatgcc tggaccacaca  
 gactgcgtga  
 601 gagaaaacag ctggtgattt atgaagagat cagcgaccct gaggaagatg  
 acgagtaact  
 661 cccctcaggg atacgacaca tgcccatgtat gagaaggcaga acgtggtgac  
 ctttcacgaa  
 721 catgggcatg gctgcccacc cctcgatc aggtgcatacg caagtg  
 10

1: **Homo sapiens folate hydrolase** [PubMed](#), [Protein](#), [Related Sequences](#), [Taxonomy](#), [OMIM](#), [LinkOut](#)  
 NM\_004 (prostate-specific membrane antigen)  
 476 1 (FOLH1), mRNA  
 LOCUS NM\_004476 2653 bp mRNA. PRI 01-  
 NOV-2000

15 DEFINITION Homo sapiens folate hydrolase (prostate-specific membrane antigen)  
 1 (FOLH1), mRNA.  
 ACCESSION NM\_004476  
 VERSION NM\_004476.1 GI:4758397  
 20 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 Euteleostomi;  
 25 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 Homo.  
 REFERENCE 1 (bases 1 to 2653)  
 AUTHORS Israeli,R.S., Powell,C.T., Fair,W.R. and Heston,W.D.  
 TITLE Molecular cloning of a complementary DNA encoding a  
 30 prostate-specific membrane antigen  
 JOURNAL Cancer Res. 53 (2), 227-230 (1993)  
 MEDLINE 93113576  
 REFERENCE 2 (bases 1 to 2653)  
 AUTHORS Rinker-Schaeffer CW, Hawkins AL, Su SL, Israeli RS,  
 35 Griffin CA,  
 Isaacs JT and Heston WD.  
 TITLE Localization and physical mapping of the prostate-  
 specific membrane  
 antigen (PSM) gene to human chromosome 11  
 40 JOURNAL Genomics 30 (1), 105-108 (1995)  
 MEDLINE 96129312  
 PUBMED 8595888  
 REFERENCE 3 (bases 1 to 2653)  
 AUTHORS O'Keefe DS, Su SL, Bacich DJ, Horiguchi Y, Luo Y,  
 45 Powell CT,  
 Zandvliet D, Russell PJ, Molloy PL, Nowak NJ, Shows  
 TB, Mullins C,  
 Vonder Haar RA, Fair WR and Heston WD.  
 TITLE Mapping, genomic organization and promoter analysis of  
 50 the human  
 prostate-specific membrane antigen gene  
 JOURNAL Biochim. Biophys. Acta 1443 (1-2), 113-127 (1998)  
 MEDLINE 99057588  
 PUBMED 9838072  
 55 REFERENCE 4 (bases 1 to 2653)

AUTHORS Maraj BH, Leek JP, Karayi M, Ali M, Lench NJ and  
 Markham AF.  
 TITLE Detailed genetic mapping around a putative prostate-  
 specific  
 5 membrane antigen locus on human chromosome 11p11.2  
 JOURNAL Cytogenet. Cell Genet. 81 (1), 3-9 (1998)  
 MEDLINE 98358137  
 PUBMED 9691167  
 COMMENT PROVISIONAL REFSEQ: This record has not yet been  
 10 subject to final NCBI review. The reference sequence was derived from  
M99487.1.  
 FEATURES Location/Qualifiers  
 15 source 1..2653  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11p11.2"  
 /sex="male"  
 20 /cell\_line="LNCaP-ATCC"  
 /cell\_type="prostate"  
 /tissue\_type="prostatic carcinoma metastatic  
 lymph node"  
 25 gene 1..2653  
 /gene="FOLH1"  
 /note="FOLH; PSM; PSMA"  
 /db\_xref="LocusID:2346"  
 /db\_xref="MIM:600934"  
 30 CDS 262..2514  
 /gene="FOLH1"  
 /note="folate hydrolase 1 (prostate-specific  
 membrane antigen)"  
 35 /codon\_start=1  
 /db\_xref="LocusID:2346"  
 /db\_xref="MIM:600934"  
 /evidence=experimental  
 /product="folate hydrolase (prostate-specific  
 40 membrane antigen) 1"  
 /protein\_id="NP\_004467.1"  
 /db\_xref="GI:4758398"  
 45 /translation="MWNLLHETDSAVATARRPRWLCA GALVLAGGFFLLGFLFGWFIK  
 SSNEATNITPKHNMKAFLDELKAENIKKFLYNFTQIPLLAGTEQNFQLAKQIQSQWKE  
 FGLDSVELAHYDVLLSYPNKTHPNYISIINEDGNEIFNTSLFEP PPPGYENVSDIVPP  
 50 FSAFSPQGMPEGDLVYVNYARTEDFFKLERDMKINC SGKIVIARYGKVFRGNKVNAQ  
 LAGAKGVILYSDPADYFAPGVKSYPDGWNLPGGGVQRGNILNLNGAGDPLTPGYPANE  
 55 YAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMGGSAPPDSSWRGSLKV PYNVGP GFT  
 GNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAEPDRYVILGGHRDSWVFGGIDPQSGA

AVVHEIVRSFGTLKKEGWRPRRTILFASWDAEFGLLGSTEWAEEENSRLLQERGVAYI  
 NADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYESWTKKSPSPEFSG  
 5 MPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYETYELVEKFY  
 DPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQEMKT  
 10 YSVSFDSLFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIGDPLGLP  
 DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAFTVQ  
 AAAETLSEVA"  
 15 misc feature 778..1029  
 /note="PA; Region: PA domain"  
 BASE COUNT 782 a 524 c 640 g 707 t  
 ORIGIN  
 1 ctcaaaaagggg gccggatttc ctctctctgg aggcatgt tgccctcttc  
 tctcgctcg 20 61 attggttcag tgcactctag aaacactgtct gtgggtggaga aactggaccc  
 caggtctgga 121 gccaattcca gcctgcaggg ctgataagcg aggcatgtt gagattgaga  
 gagactttac 181 cccgcgtgg tgggtggagg gcgcgcgtta gagcagcgc acaggcgcgg  
 25 gtcccccggag 241 gccggctctg ctgcgcgcga gatgtggat ctcccttcacg aaaccgactc  
 ggctgtggcc 301 accgcgcgcgc gcccgcgtg gctgtgcgtt ggggcgcgtgg tgctggcggg  
 tggcttctt 361 ctccctcggtt tcctcttcgg gtggttata aaatcctcca atgaagctac  
 taacattact 421 ccaaagcata atatgaaagc atttttggat gaattgaaag ctgagaacat  
 caagaagtcc 481 ttatataatt ttacacagat accacattta gcaggaacag aacaaaactt  
 35 tcagcttgca 541 aagcaaattt aatcccagtg gaaaaggat ggcctggatt ctgttgagct  
 agcacattat 601 gatgtcctgt tgtcctaccc aaataagact catcccaact acatctcaat  
 aatataatgaa 40 661 gatggaaatg agatttcaa cacatcatta tttgaaccac ctccctccagg  
 atatgaaaat 721 gtttcggata ttgtaccacc tttcagtgtct ttctctcctc aaggaatgcc  
 agagggcgat 781 ctatgtatg ttaactatgc acgaactgaa gacttcttta aattggaaacg  
 ggacatgaaa 841 atcaattgtct ctggaaaat tgtaattgcc agatatggaa aagttttcag  
 agggaaaatg 901 gttaaaaatg cccagctggc agggggccaaa ggagtcattc tctactccga  
 ccctgctgac 50 961 tactttgtct ctggggtgaa gtcctatcca gatggttgga atcttcctgg  
 aggtgggtgtc 1021 cagcgtggaa atatccaaa tctgaatggt gcaggagacc ctctcacacc  
 aggttaccca 1081 gcaaattgaaat atgcttatacg gcgtggaaatt gcagaggctg ttggcttcc  
 55 aagtattccat 1141 gttcatccaa ttggatacta tgatgcacag aagctccttag aaaaaatggg  
 tggctcagca

1201 ccaccagata gcagctggag aggaagtctc aaagtgcctc acaatgttgg  
 acctggctt  
 1261 actggaaact tttctacaca aaaagtcaag atgcacatcc actctaccaa  
 5 tgaagtgaca  
 1321 agaatttaca atgtgatagg tactctcaga ggagcagtgg aaccagacag  
 atatgtcatt  
 1381 ctgggaggtc accgggactc atgggtgtt ggtggattt accctcagag  
 tggagcagct  
 1441 gtgttcatg aaattgtgag gagctttgga acactgaaaa aggaagggtg  
 10 gagacccataga  
 1501 agaacaattt tgtttgcaag ctggatgca gaagaatttg gtcttcttgg  
 ttctactgag  
 1561 tggcagagg agaattcaag actcctcaa gagcgtggcg tggcttatat  
 15 taatgctgac  
 1621 tcatactatag aaggaaacta cactctgaga gttgatttga caccgctgat  
 gtacagctt  
 1681 gtacacaacc taacaaaaga gctgaaaagc cctgatgaag gctttgaagg  
 20 caaatcttctt  
 1741 tatgaaaagtt ggactaaaaa aagtccttcc ccagagttca gtggcatgcc  
 caggataagc  
 1801 aaattggat ctggaaatga ttttgaggtg ttcttccaac gacttggaaat  
 tgcttcaggc  
 1861 agagcaggt atactaaaaa ttggaaaca aacaaattca gcggctatcc  
 actgtatcac  
 25 1921 agtgtctatg aaacatatga gttggggaa aagttttatg atccaatgtt  
 taaatatcac  
 1981 ctcaactgtgg cccaggttcg aggagggatg gtgtttgagc tagccaaatcc  
 catatgctc  
 2041 cctttgatt gtcgagatta tgctgttagtt ttaagaaaat atgctgacaa  
 30 aatctacagt  
 2101 atttctatga aacatccaca ggaaatgaag acatacagtg tatcatttga  
 ttcactttt  
 2161 tctgcagtaa agaattttac agaaatttgc tccaaaggta gtgagagact  
 ccaggactt  
 35 2221 gacaaaagca acccaatagt attaagaatg atgaatgatc aactcatgtt  
 tctggaaaga  
 2281 gcattttattg atccattagg gttaccagac aggcctttt ataggcatgt  
 catctatgtc  
 2341 ccaaggcagcc acaacaagta tgcaaaaaag tcattccag gaattttatga  
 40 tgctctgttt  
 2401 gatattgaaa gcaaaagtggc cccttccaag gcctggggag aagtgaagag  
 acagatttat  
 2461 gtgcagcct tcacagtgc ggcagctgc gagactttga gtgaagtagc  
 ctaagaggat  
 2521 tcttttagaga atccgtattt aatttggatg gtatgtcact cagaaagaat  
 45 cgtaatgggt  
 2581 atattgataa attttaaaat tggatattt gaaataaaat tgaatattat  
 atataaaaaaa  
 2641 aaaaaaaaaaaa aaa

U20093. Human melanocyte-... [gi:1142634]  
 LOCUS HSPMEL17S2 2817 bp DNA PRI 20-  
 NOV-1996  
 DEFINITION Human melanocyte-specific (pmel 17) gene, exons 2-5,  
 5 and complete  
 cds.  
 ACCESSION U20093  
 VERSION U20093.1 GI:1142634  
 KEYWORDS  
 10 SEGMENT 2 of 2  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 15 Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 Homo.  
 REFERENCE 1 (bases 1 to 2817)  
 AUTHORS Kim, K.K., Youn, B.S., Heng, H.H., Shi, X.M., Tsui, L.C.,  
 20 Lee, Z.H.,  
 Pickard, R.T. and Kwon, B.S.  
 TITLE Genomic organization and FISH mapping of human Pmel  
 17, the  
 putative silver locus  
 25 JOURNAL Pigment Cell Res. 9 (1), 42-48 (1996)  
 MEDLINE 96314705  
 REFERENCE 2 (bases 1 to 2817)  
 AUTHORS Kwon, B.S.  
 TITLE Direct Submission  
 30 JOURNAL Submitted (05-JAN-1995) Indiana University School of  
 Medicine,  
 Microbiology and Immunology, 635 Barnhill Drive,  
 Indianapolis, IN  
 46202, USA  
 FEATURES Location/Qualifiers  
 35 source 1..2817  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /map="12pter-q21"  
 40 gene /clone\_lib="lambda FIX II from Stratagene"  
 join(U19491.1:1111..3109,1..2722)  
 /gene="pmel 17"  
 CDS  
 45 join(U19491.1:1111..1186,67..459,853..1014,1266..2396,  
 2499..2722)  
 /gene="pmel 17"  
 /codon\_start=1  
 /product="Pmel 17"  
 /protein\_id="AAB19181.1"  
 50 /db\_xref="GI:1142636"  
 /translation="MDLVLKRCLLHLAVIGALLAVGATKVRNQDWLGVSRLRTKAWNRLQYPEW  
 TEAQLDCWRGGQVSLKVSNDGPTLIGANASFSIALNFPGSQKVLPDGQVIWVNNTIINGSQVWGG  
 QPVYPQETDDACIFPDGGPCPSGSWSQKRSFVYVWKTWQYWQVLGGPVSGLSIGTGRAMLGHTM  
 55 EVTYYHRRGSRSYVPLAHSSAFTITDQVPGFVSVSQLRALDGGNKHFLRNQPLTFALQLHDPSGY  
 LAEADLSYTWDFFGDSSGTLISRAPVVTHTYLEPGPVTAQVVLQAAIPLTSCGSSPVPGTTDGHRT  
 AEAPNTTAGQVPTTEVVGTTPGQAPTAEPGTTSVQVPTTEVISTAPVQMPTESTGMTPEKVPVS  
 EVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTTTEWVETTARELPIPEPEGPDASSIMSTE"

SITGSLGPLLDGTATLRLVKRQVPLDCVLYRYGSFSVTLDIVQGIESAEILQAVPSGEGDAFELTV  
 SCQGGGLPKEACMEIISSPGCQPPAQRLCPVLPSPACQLVLHQILKGGSGTYCLNVSLADTNSLAVV  
 STQLIMPQEAQLGQVPLIVGILLVLMAVVLAISLYRRRLMKQDFSVPQLPHSSSHWLRLPRIFCS  
 CPIGENSPLLSGQQV"  
 5           exon           67..459  
 /gene="pmel 17"  
 /number=2  
 exon           853..1014  
 /gene="pmel 17"  
 /number=3  
 10           exon           1266..2396  
 /gene="pmel 17"  
 /number=4  
 exon           2499..>2722  
 /gene="pmel 17"  
 /number=5  
 BASE COUNT    627 a    772 c    739 g    679 t  
 ORIGIN  
 20           1 gtgctaaaaa gatgccttct tcatttggct gtgataggtg ctttgtggct  
 gtgggggctta  
 61           61 caaaaagtacc cagaaaccag gactggcttg gtgtctcaag gcaactcaga  
 accaaaggcct  
 121           121 ggaacaggca gctgtatcca gagtggacag aagccagag acttgactgc  
 tggagaggtg  
 25           181 gtcaagtgtc cctcaaggc agtaatgatg ggcctacact gattggtgca  
 aatgcctcct  
 241           241 tctctattgc cttgaacttc cctggaagcc aaaaggtatt gccagatggg  
 caggttatct  
 30           301 gggtcaacaa taccatcatc aatgggagcc aggtgtgggg aggacagcca  
 gtgtatcccc  
 361           361 aggaaactga cgatgcctgc atcttccctg atggtgacc ttgcccatt  
 ggctcttgggt  
 421           421 ctcagaagag aagctttgtt tatgtctgga agacctgggg tgagggactc  
 ccttctcagc  
 481           481 statcatcca cacttgtgtt tacttcttc tacctgatca ctttttttt  
 ggcccccct  
 541           541 tccacctaa cttctgtat tttctctaatt cttcattttc ctcttagatc  
 ttttctcttt  
 601           601 cttagcacct agcccccttc aagctctatc ataattttt ctggcaactc  
 40           ttggcctcaa  
 661           661 ttgttagtcct accccatgga atgcctcatt aggacccctt ccctgtcccc  
 ccatatcaca  
 721           721 gccttccaaa caccctcaga agtaatcata cttcctgacc tcccatctcc  
 agtggcgttt  
 781           781 cgaaggctgt ccctcagtcc ctttgcacca gtaatcttt ctccatttgct  
 tttcattcca  
 841           841 aaaatgcttc aggccaatac tggcaagttc tagggggccc agtgtctggg  
 ctgagcattg  
 901           901 ggacaggcag ggcaatgctg ggcacacaca ccatggaagt gactgtctac  
 catgcgggg  
 961           961 gatcccggag ctatgtgcct cttgctcatt ccagctcagc cttcaccatt  
 actgtaagg  
 1021           1021 gttcaggaag ggcaaggcca gttgttaggg aaagagaagg cagggaggct  
 tggatggact  
 1081           1081 gcaaaggaga aaggtgaaat gctgtcaaa cttaaagtag aagggccagg  
 aagacctagg  
 1141           1141 cagagaaatg tgaggcttag tgccagtgaa gggccagcca gtcagcttgg  
 agttggaggg

	1201	tgtggctgtg	aaaggagaag	ctgtggctca	ggcctggttc	tcaccttttc
	1261	tggctccat	cccagaccag	gtgccttct	ccgtgagcgt	gtcccagttg
5	atggagggaa	1321	caagcacttc	ctgagaaaatc	agcctctgac	ctttgccctc
	accccagtgg	1381	ctatctggct	gaagctgacc	tctcctacac	ctgggacttt
	gtggaaccct	1441	gatctctcg	gcacctgtgg	tcactcatac	ttacctggag
10	tcactgccc	1501	gttggtcctg	caggctgcca	ttcctctcac	ctcctgtggc
	ttccaggcac	1561	cacagatggg	cacaggccaa	ctgcagaggc	ccctaaacacc
15	aagtgcctac	1621	tacagaagtt	gtgggtacta	cacctggtca	ggcgccaaact
	ctggaaaccac	1681	atctgtgcag	gtgccaacca	ctgaagtcat	aagcactgca
	tgccaaactgc	1741	agagagcaca	ggtatgacac	ctgagaaggt	gccagttca
20	gtaccacact	1801	ggcagagatg	tcaactccag	aggctacagg	tatgacaccc
	caatttgtgt	1861	gtttctgga	accacagctg	cacaggtaac	aactacagag
25	ccacagctag	1921	agagctacct	atccctgagc	ctgaagggtcc	agatgccagc
	ctacggaaag	1981	tattacaggt	tccttggcc	ccctgctgga	tggtacagcc
30	tggtaagag	2041	acaagtcccc	ctggattgtg	ttctgtatcg	atatggttcc
	cccttggacat	2101	tgtccagggt	attgaaagtg	ccgagatcct	gcaggctgtg
	agggggatgc	2161	attttagctg	actgtgtcct	gccaaggcgg	gctgccaag
35	tggagatctc	2221	atcgccaggg	tgccagcccc	ctgcccagcg	gctgtgccag
	ccagcccgac	2281	ctgccagctg	gttctgcacc	agataactgaa	gggtggctcg
	gcctcaatgt	2341	gtctctggct	gataccaaca	gcctggcagt	ggtcagcacc
40	tgccttggtag	2401	gtccttggac	agagactaag	tgaggaggga	cagtttatca
	ggcaagcagc	2461	agacatgagt	gaagcagtgc	ctgggattct	tctcacaggt
45	gccttggca	2521	gttcccgctg	atcggtggca	tcttgctggt	caagaagcag
	catctctgat	2581	atataggcgc	agacttatga	gttgatggct	ggggacagct
	atagcagcag	2641	tcactggctg	cgtctacccc	gtggtccttg	atgggtgaga
50	atagccccct	2701	cctcagtggg	cagcagggtct	ctcttgcacc	atatgatgct
	tggagttgac	2761	agaaacaccc	atatttcccc	cagtcttccc	gtgattttcc
55	aaataaaa	//				

NM\_001648. Homo sapiens kall...[gi:4502172]

LOCUS NM\_001648 1466 bp mRNA PRI  
31-OCT-2000  
DEFINITION Homo sapiens kallikrein 3, (prostate specific antigen)  
(KLK3),  
5 mRNA.  
ACCESSION NM\_001648  
VERSION NM\_001648.1 GI:4502172  
KEYWORDS  
10 SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata;  
Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini;  
Hominidae; Homo.  
15 REFERENCE 1 (bases 1 to 1466)  
AUTHORS Lundwall,A. and Lilja,H.  
TITLE Molecular cloning of human prostate specific antigen  
cDNA  
JOURNAL FEBS Lett. 214 (2), 317-322 (1987)  
20 MEDLINE 87190978  
REFERENCE 2 (bases 1 to 1466)  
AUTHORS Sutherland GR, Baker E, Hyland VJ, Callen DF, Close JA,  
Tregear GW,  
Evans BA and Richards RI.  
25 TITLE Human prostate-specific antigen (APS) is a member of  
the glandular  
kallikrein gene family at 19q13  
JOURNAL Cytogenet. Cell Genet. 48 (4), 205-207 (1988)  
MEDLINE 89250658  
30 PUBMED 2470553  
REFERENCE 3 (bases 1 to 1466)  
AUTHORS Riegman PH, Klaassen P, van der Korput JA, Romijn JC  
and Trapman J.  
TITLE Molecular cloning and characterization of novel  
35 prostate antigen  
cDNA's  
JOURNAL Biochem. Biophys. Res. Commun. 155 (1), 181-188 (1988)  
MEDLINE 88326297  
PUBMED 2458104  
40 REFERENCE 4 (bases 1 to 1466)  
AUTHORS Schulz P, Stucka R, Feldmann H, Combriato G, Klobbeck HG  
and Fittler F.  
TITLE Sequence of a cDNA clone encompassing the complete  
mature human  
45 prostate specific antigen (PSA) and an unspliced  
leader sequence  
JOURNAL Nucleic Acids Res. 16 (13), 6226 (1988)  
MEDLINE 88289366  
PUBMED 2456523  
50 REFERENCE 5 (bases 1 to 1466)  
AUTHORS Riegman PH, Vlietstra RJ, van der Korput JA, Romijn JC  
and Trapman J.  
TITLE Characterization of the prostate-specific antigen gene:  
a novel  
55 human kallikrein-like gene  
JOURNAL Biochem. Biophys. Res. Commun. 159 (1), 95-102 (1989)  
MEDLINE 89165891  
PUBMED 2466464

REFERENCE 6 (bases 1 to 1466)  
AUTHORS Henttu P and Vihko P.  
TITLE cDNA coding for the entire human prostate specific  
antigen shows  
5 high homologies to the human tissue kallikrein  
genes  
JOURNAL Biochem. Biophys. Res. Commun. 160 (2), 903-910 (1989)  
MEDLINE 89246551  
PUBMED 2470373  
10 REFERENCE 7 (bases 1 to 1466)  
AUTHORS Klobbeck HG, Combriato G, Schulz P, Arbusow V and  
Fittler F.  
TITLE Genomic sequence of human prostate specific antigen  
(PSA)  
15 JOURNAL Nucleic Acids Res. 17 (10), 3981 (1989)  
MEDLINE 89282407  
PUBMED 2471958  
REFERENCE 8 (bases 1 to 1466)  
AUTHORS Lundwall A.  
20 TITLE Characterization of the gene for prostate-specific  
antigen, a human  
glandular kallikrein  
JOURNAL Biochem. Biophys. Res. Commun. 161 (3), 1151-1159  
(1989)  
25 MEDLINE 89302090  
PUBMED 2472789  
COMMENT PROVISIONAL REFSEQ: This record has not yet been  
subject to final  
NCBI review. The reference sequence was derived  
30 from X05332.1.  
FEATURES Location/Qualifiers  
source 1..1466  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
35 /chromosome="19"  
/map="19q13"  
/clone="lambda HPSA-1"  
/tissue\_type="prostate"  
/clone\_lib="(lambda)gt11"  
40 gene 1..1466  
/gene="KLK3"  
/note="PSA; APS"  
/db\_xref="LocusID:354"  
/db\_xref="MIM:176820"  
45 sig\_peptide 44..94  
CDS 44..829  
/gene="KLK3"  
/EC\_number="3.4.21.77"  
/note="antigen, prostate specific"  
50 /codon\_start=1  
/db\_xref="LocusID:354"  
/db\_xref="MIM:176820"  
/product="kallikrein 3, (prostate specific  
antigen)"  
55 /protein\_id="NP\_001639.1"  
/db\_xref="GI:4502173"

```

/translation="MWVPVVFLTLSVTWIGAAPLILSRIVGGWECEKHSQPWQVLVAS
RGRAVCGGVLVHPQWLTAACIRNKSILLGRHSLFHPEDTGQVFQVSHSFPHPLYDMSLLKNRF
LPGDDSSHDLMLRLSEPAELTDAVKVMDLPTQEALGTTCYASGWGSIEPEEFLTPKKLQCVDL
5 HVISNDVCAQVHPQKVTKFMLCAGRWTGGKSTCSGDGGPLVCNGVLQGITSWGSEPCALPERPSL
YTKVVHYRKWIKDTIVANP"
    variation      91
        /allele="C"
        /allele="T"
        /db_xref="dbSNP:1135765"
10    variation      91
        /allele="C"
        /allele="T"
        /db_xref="dbSNP:11573"
15    misc_feature  95..115
        /note="propeptide (AA 1-7)"
    variation      97
        /allele="G"
        /allele="A"
20    misc_feature  113..802
        /db_xref="dbSNP:1135766"
        /note="Tryp_SPC; Region: Trypsin-like serine
protease"
25    misc_feature  116..802
        /note="trypsin; Region: Trypsin"
    mat_peptide    116..826
        /product="kallikrein 3, (prostate specific
antigen)"
30    variation      280
        /allele="C"
        /allele="T"
        /db_xref="dbSNP:1058072"
    variation      280
        /allele="C"
        /allele="T"
        /db_xref="dbSNP:12946"
35    variation      526
        /allele="A"
        /allele="G"
        /db_xref="dbSNP:1803130"
40    variation      643
        /allele="A"
        /allele="G"
        /db_xref="dbSNP:1803129"
45    variation      788
        /note="WARNING: map location ambiguous"
        /allele="T"
        /allele="C"
        /db_xref="dbSNP:1802720"
50    variation      844
        /allele="T"
        /allele="C"
        /db_xref="dbSNP:1058205"
55    variation      894
        /allele="A"
        /allele="G"
        /db_xref="dbSNP:1058274"
    variation      906

```

```

5           variation   /allele="T"
             /allele="A"
             /db_xref="dbSNP:1803132"
918
10          variation  /allele="T"
             /allele="C"
             /db_xref="dbSNP:1802722"
953
15          variation  /allele="A"
             /allele="G"
             /db_xref="dbSNP:1802721"
1017
20          variation  /allele="T"
             /allele="C"
             /db_xref="dbSNP:1803138"
1074
25          variation  /allele="C"
             /allele="T"
             /db_xref="dbSNP:12101"
1107
30          variation  /allele="A"
             /allele="G"
             /db_xref="dbSNP:1803131"
1107
35          variation  /allele="G"
             /allele="A"
             /db_xref="dbSNP:6998"
1211
40          variation  /allele="C"
             /allele="T"
             /db_xref="dbSNP:1803136"
1229
45          variation  /allele="A"
             /allele="G"
             /db_xref="dbSNP:1803135"
1249
50          variation  /allele="T"
             /allele="C"
             /db_xref="dbSNP:1803134"
1273
55          variation  /allele="A"
             /allele="G"
             /db_xref="dbSNP:12040"
1326
           misc_feature /db_xref="dbSNP:1803133"
1344
           variation   /allele="A"
             /allele="G"
             /db_xref="dbSNP:12863"
1363
           variation   /allele="A"
             /allele="G"
             /db_xref="dbSNP:1803137"
1445..1450
           polyA_site   /note="put.polyA signal"
1466

```

```

/note="polyA site"
BASE COUNT      338 a      382 c      422 g      324 t
ORIGIN
      1 agccccaaagc ttaccacctg caccggaga gctgtgtgtc accatgtggg
      5 tcccggttgt
      61 cttcctcacc ctgtccgtga cgtggattgg tgctgcaccc ctcatcctgt
      ctcggattgt
      121 gggaggctgg gagtgcgaga agcattccca accctggcag gtgcttgtgg
      10 cctctgtgg
      181 cagggcagtc tgcggcggtg ttctggtgca cccccagtgg gtcctcacag
      ctgcccactg
      241 catcaggaac aaaagcgtga tcttgctggg tcggcacagc ctgtttcatc
      ctgaagacac
      301 aggccaggtt tttcaggtca gccacagtt cccacacccg ctctacgata
      15 tgagcctcct
      361 gaagaatcga ttccctcaggc caggtgatga ctccagccac gacctcatgc
      tgctccgcct
      421 gtcagagcct gccgagctca cggatgctgt gaaggtcatg gacctgcca
      cccaggagcc
      481 agcaactgggg accacctgct acgcctcagg ctggggcagc attgaaccag
      aggagttctt
      541 gaccccaaag aaacttcagt gtgtggaccc ccatgttatt tccaatgacg
      tgtgtgcgca
      601 agttcacccct cagaagggtga ccaagttcat gctgtgtgtc ggacgctgg
      20 cagggggcaa
      661 aagcacctgc tcgggtgatt ctgggggccc acttgtctgt aatggtgtc
      ttcaaggtat
      721 cacgtcatgg ggcagtgaac catgtccct gcccggaaagg cttccctgt
      acaccaaggt
      781 ggtgcattac cggaaagtggc tcaaggacac catcggtggcc aacccctgag
      30 cacccttatac
      841 aacccctat ttagtagaaac ttggAACCTT ggaaatgacc aggccaagac
      tcaaggcctcc
      901 ccagtttac tgacccctgt ccttaggtgt gaggtccagg gttgttagga
      35 aaagaaatca
      961 gcagacacag gtgttagacca gagtgtttct taaatgggtt aattttgtcc
      tcttgtgtc
      1021 ctggggaaata ctggccatgc ctggagacat atcactcaat ttctctgagg
      acacagatag
      1081 gatgggggtgt ctgtgttatt tgggggtac agagatgaaa gaggggtggg
      40 atccacactg
      1141 agagagtggc gagtgacatg tgctggacac tggccatgaa gcactgacca
      gaagctggag
      1201 gcacaacgca ccagacactc acagcaagg tggagctgaa aacataaccc
      45 actctgtcct
      1261 ggaggcactg ggaaggctag agaaggctgt gagccaagga gggagggtct
      tcccttggca
      1321 tgggatgggg atgaagtaag gagagggact ggacccctg gaagctgatt
      50 cactatgggg
      1381 ggaggtgtat tgaagtcctc cagacaaccc tcagatttga tgatttccta
      gtagaactca
      1441 cagaaataaa gagctgttat actgtg
      //

```

55 U87459. Human autoimmunog...[gi:1890098]  
 LOCUS HSU87459 752 bp mRNA PRI  
 22-DEC-1999

DEFINITION Human autoimmunogenic cancer/testis antigen NY-ESO-1 mRNA, complete cds.  
 5 ACCESSION U87459  
 VERSION U87459.1 GI:1890098  
 KEYWORDS .  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata;  
 10 Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 752)  
 AUTHORS Chen, Y.T., Scanlan, M.J., Sahin, U., Tureci, O.,  
 15 Gure, A.O., Tsang, S., Williamson, B., Stockert, E., Pfreundschuh, M. and Old, L.J.  
 TITLE A testicular antigen aberrantly expressed in human cancers detected  
 by autologous antibody screening  
 20 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (5), 1914-1918 (1997)  
 MEDLINE 97203161  
 PUBMED 9050879  
 REFERENCE 2 (bases 1 to 752)  
 AUTHORS Chen, Y.-T.  
 25 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) Ludwig Institute for Cancer Research, New York Branch, 1275 York Avenue, New York, NY 10021, USA  
 30 FEATURES Location/Qualifiers  
 source 1..752  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="esophageal squamous cell  
 35 carcinoma expression cDNA library"  
 CDS 54..596  
 /codon\_start=1  
 /product="autoimmunogenic cancer/testis  
 40 antigen NY-ESO-1"  
 /protein\_id="AAB49693.1"  
 /db\_xref="GI:1890099"  
 /translation="MQAEGRGTGGSTGDADGPGGPGIPDGPAGNAGGPGEAG  
 45 ATGGGRGPRGAGAARASGPAGGAPRGPHGGAASGLNGCCRCGARGPESRLLEF  
 YLAMPFATPMEALARRSLAQDAPPVPVGVLKЕFTVSGNILTIRLTAADH  
 RQLQLSISSCLQQLSLLMWITQCFLPVFLAQPPSGQRR"  
 BASE COUNT 126 a 230 c 256 g 140 t  
 50 ORIGIN  
 1 atcctcggtt gccctgacct tctctctgag agccggcag aggctccgg  
 gccatgcagg  
 61 ccgaaggccg gggcacagg ggttcgacgg gcgatgctga tggcccagga  
 ggcctggca  
 121 ttccctgatgg cccaggggc aatgctggcg gcccaggaga ggcgggtgcc  
 55 acggggcgca  
 181 gaggtccccg gggcgccagg gcagcaagg cctcggggcc gggaggaggc  
 gccccggggc

```

      241 gtccgcattt cggcgccgt tcagggctga atggatgctg cagatgcggg
      301 cggagagccg cctgctttag ttctacctcg ccatgcctt cgcgacaccc
      361 agctggcccg caggagcctg gcccaggatg ccccaccgct tcccgtgcca
      421 tgaaggagtt cactgtgtcc ggcaacatac tgactatccg actgactgt
      481 gccaactgca gcttcattc agtcctgtc tccagcagct ttccctgttg
      541 cgcagtgcctt tctgcccgtg tttttggctc agcctccctc agggcagagg
      601 agcctggcgc cccttcctag gtcatgcctc ctccccttagg gaatggtccc
      661 gccagttcat tggggggcc tgattgtttg tcgctggagg aggacggcgtt
      721 ttctgttagaa aataaaaactg agctacgaaa aa
      //

20   CAA11116. LAGE-1a protein [...[gi:3255959]
LOCUS          CAA11116          180 aa
23-JUN-1998
DEFINITION LAGE-1a protein [Homo sapiens].
ACCESSION      CAA11116
25   PID            g3255959
VERSION        CAA11116.1 GI:3255959
DBSOURCE       embl locus HOS223093, accession AJ223093.1
KEYWORDS
SOURCE          human.
30   ORGANISM      Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata;
                  Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini;
                  Hominidae; Homo.
35   REFERENCE     1 (residues 1 to 180)
AUTHORS        Lethe,B., Lucas,S., Michaux,L., De Smet,C.,
                  Godelaine,D.,
                  Serrano,A., De Plaen,E. and Boon,T.
40   TITLE          LAGE-1, a new gene with tumor specificity
JOURNAL        Int. J. Cancer 76 (6), 903-908 (1998)
MEDLINE        98289662
REFERENCE     2 (residues 1 to 180)
45   AUTHORS        Lethe,B.G.
TITLE          Direct Submission
JOURNAL        Submitted (08-JAN-1998) Lethe B.G., Brussels Branch,
                  Ludwig
                  Institute for Cancer Research, 74 Avenue
                  Hippocrate, B - 1200 -
                  Bruxelles, BELGIUM
50   COMMENT        Related sequences: AJ223040, AJ223041 and AJ003149.
FEATURES       Location/Qualifiers
source          1..180
                  /organism="Homo sapiens"
                  /isolate="individual LB33"
                  /db_xref="taxon:9606"
                  /chromosome="X"
                  /map="q28"
                  /clone="c2RB"
55

```

```

      /sex="Female"
      /cell_line="LB33-MELA"
      /cell_type="melanoma"
      /clone_lib="B6"
5       Protein      1..180
      /product="LAGE-1a protein"
      CDS          1..180
      /gene="LAGE-1"
      /db_xref="SPTREMBL:O75637"
10      /coded_by="join(AJ223093.1:1231..1499,
                  AJ223093.1:2115..2249,AJ223093.1:2479..2617)"
```

ORIGIN

```

15      1 mqaegrgrtgg stgdadgpagg pgipdgpggn agggpeagat ggrgprgaga
      arasgprgga
      61 prgphggaas aqdgrcpcga rrpdsrllel hitmpfsspm eaelvrrils
      rdaaplprpg
      121 avlkdftvsg nllfirltaa dhrqlqlsis sclqqlsllm witqcfvpvf
      laqapsgqrr
20      181
      //
```

**CAA11117. LAGE-1b protein [...[gi:3255960]**

LOCUS CAA11117 210 aa PRI

25 23-JUN-1998

DEFINITION LAGE-1b protein [Homo sapiens].

ACCESSION CAA11117

PID g3255960

VERSION CAA11117.1 GI:3255960

30 DBSOURCE embl locus HOS223093, accession AJ223093.1

KEYWORDS .

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata;

35 Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (residues 1 to 210)

AUTHORS Lethe,B., Lucas,S., Michaux,L., De Smet,C.,
40 Godelaine,D.,
 Serrano,A., De Plaen,E. and Boon,T.

TITLE LAGE-1, a new gene with tumor specificity

JOURNAL Int. J. Cancer 76 (6), 903-908 (1998)

MEDLINE 98289662

45 REFERENCE 2 (residues 1 to 210)

AUTHORS Lethe,B.G.

TITLE Direct Submission

JOURNAL Submitted (08-JAN-1998) Lethe B.G., Brussels Branch,
50 Ludwig
 Institute for Cancer Research, 74 Avenue
 Hippocrate, B - 1200 -
 Bruxelles, BELGIUM

COMMENT Related sequences: AJ223040, AJ223041 and AJ003149.

FEATURES Location/Qualifiers

55 source 1..210
 /organism="Homo sapiens"
 /isolate="individual LB33"
 /db\_xref="taxon:9606"

```

5           /chromosome="X"
           /map="q28"
           /clone="c2RB"
           /sex="Female"
           /cell_line="LB33-MELA"
           /cell_type="melanoma"
           /clone_lib="B6"
           1..210
           Protein      /product="LAGE-1b protein"
10          CDS          1..210
           /gene="LAGE-1"
           /db_xref="SPTREMBL:O75638"
           /coded_by="join(AJ223093.1:1231..1499,
AJ223093.1:2115..2478)"
           /note="alternate splicing at exon 2 - intron
15          2 boundary"
           ORIGIN
           1 mqaegrktgg stgdadgpgg pgipdgpggn aggpgeagat ggrgprgaga
           arasgprgga
20          61 prgphggaas aqdgrcpcga rrpdsrllel hitmpfsspm eaelvrrils
           rdaaplprpg
           121 avlkdftvsg nllfmsvwdq dregagrmrv vgwglgsasp eggkardlrt
           pkhkvseqrp
           181 gtpgppppeg aqgdgcrgva fnvmfsaphi
25          //

```

```

30          M77481. Human antigen (MA...[gi:416114]
LOCUS          HUMMAG1A      2420 bp      DNA          PRI
15-NOV-1993
DEFINITION  Human antigen (MAGE-1) gene, complete cds.
ACCESSION  M77481
35          VERSION      M77481.1  GI:416114
KEYWORDS  antigen.
SOURCE          Homo sapiens (individual_isolate patient MZ2)
melanoma abdominal
                           metastasis of melanoma DNA.
40          ORGANISM  Homo sapiens
                           Eukaryota;      Metazoa;      Chordata;      Craniata;
Vertebrata;  Euteleostomi;
                           Mammalia;      Eutheria;      Primates;      Catarrhini;
Hominidae;  Homo.
45          REFERENCE  1 (bases 785 to 1286)
AUTHORS        van der Bruggen,P., Traversari,C., Chomez,P.,
Lurquin,C., De
                           Plaen,E., Van den Eynde,B., Knuth,A. and Boon,T.
50          TITLE        A gene encoding an antigen recognized by cytolytic T
lymphocytes on
                           a human melanoma
JOURNAL        Science 254, 1643-1647 (1991)
MEDLINE        92086861
55          REFERENCE  2 (bases 1 to 2420)
AUTHORS        van der Bruggen P.
TITLE          Direct Submission
JOURNAL        Submitted (05-FEB-1992) Pierre van der Bruggen, Ludwig
Institute

```

for Cancer Research, Brussels Branch, Avenue  
Hippocrate, 74, UCL  
7459, Brussels, B-1200, Belgium

COMMENT On Nov 15, 1993 this sequence version replaced  
5 gi:187294.

FEATURES Location/Qualifiers

source 1..2420  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="X"  
/sex="female"  
/cell\_line="MZ2-MEL.43"

10 intron 1..412  
/partial  
/gene="MAGE-1"  
/number=1

15 exon 413..485  
/gene="MAGE-1"  
/number=2

20 mRNA join(413..485,561..2111)  
/partial  
/gene="MAGE-1"  
intrton 486..560  
/gene="MAGE-1"  
/number=2

25 exon 561..2111  
/gene="MAGE-1"  
/number=3

30 gene 626..1555  
/gene="MAGE-1"  
CDS 626..1555  
/gene="MAGE-1"  
/codon\_start=1  
/product="MAGE-1"  
35 /protein\_id="AAA03229.1"  
/db\_xref="GI:416115"

/translation="MSLEQRSLHCKPEEALEAQQEALGLVCVQAATSSSSPL  
VLGTLLEEVPTAGSTDPPQSPQGASAFPTTINFTRQRQPSEGSSREEGPST  
40 SCILESFLFRAVITKKVADLVGFLLKYRAREFVTKAEMLESVIKNYKHCPE  
IFGKASESLQLVFGIDVKRADPTGHSYVLVTCLGLSYDGLLGDNQIMPKTGF  
LIIIVLVMIAMEGGHAAPEEEIWEELSVMEVYDGREHSAYGEPRKLLTQDLVQE  
KYLEYRQVPDSDPARYEFLWGPRALAETSYVKLEYVIKVSARVFFFPSLR  
EAALREEEEGV"

45 polyA\_signal 2095..2100  
/gene="MAGE-1"  
intron 2112..2420  
/partial  
/gene="MAGE-1"  
/number=3

50 BASE COUNT 562 a 582 c 677 g 599 t  
ORIGIN  
1 ggatccaggc cctgccagga aaaatataag ggccctgcgt gagaacagag  
ggggtcatcc  
55 61 actgcatgag agtggggatg tcacagagtc cagcccaccc tcctggtagc  
actgagaagc  
121 cagggctgtg cttgcggtct gcaccctgag ggcccgtgga ttccctttcc  
tggagctcca

	181	ggaaccaggc	agtgaggcct	tggtctgaga	cagtatcctc	aggtcacaga
	gcagaggatg					
	241	cacagggtgt	gccagcagtg	aatgtttgcc	ctgaatgcac	accaagggcc
5	ccacactgcca					
	301	caggacacat	aggactccac	agagtctggc	ctcacctccc	tactgtcagt
	cctgtagaat					
	361	cgacctctgc	tggccggctg	taccctgagt	accctctcac	ttcctccttc
	aggttttcag					
10	421	ggcacaggcc	aacccagagg	acaggattcc	ctggagggcc	cagaggagca
	ccaaggagaa					
	481	gatctgttaag	taggcctttg	ttagagtctc	caaggttcag	ttctcagctg
	aggccctctca					
	541	cacactccct	ctctccccag	gcctgtgggt	cttcattgcc	cagctcctgc
15	ccacactcct					
	601	gcctgctgcc	ctgacgagag	tcatcatgtc	tcttgagcag	aggagtctgc
	actgcaagcc					
	661	tgaggaagcc	cttgaggccc	aacaagaggc	cctgggcctg	gtgtgtgtgc
	aggctccac					
20	721	ctccctctcc	tctcctctgg	tcctgggcac	cctggaggag	gtgcccactg
	ctgggtcaac					
	781	agatcctccc	cagagtctc	agggagcctc	cgccttccc	actaccatca
	acttcactcg					
	841	acagaggcaa	cccagtgagg	gttccagcag	ccgtgaagag	gaggggccaa
	gcacctcttg					
	901	tatcctggag	tccttggtcc	gagcagtaat	caactaagaag	gtggctgatt
	tggttgggtt					
	961	tctgtccctc	aaatategag	ccagggagcc	agtcacaaag	gcagaaatgc
	tggagagtgt					
25	1021	catcaaaaaat	tacaagact	gttttcctga	gatcttcggc	aaagcctctg
	agtcccttgc					
	1081	gctggcttt	ggcattgacg	tgaaggaagc	agaccccacc	ggccactcct
	atgtccttgc					
	1141	cacctgccta	ggtctctct	atgatggcct	gctgggtgat	aatcagatca
	tgcccaagac					
30	1201	aggcttcctg	ataattgtcc	tggtcatgat	tgcaatggag	ggcgccatg
	ctcctgagga					
	1261	ggaaatctgg	gaggagctga	gtgtgatgga	ggtgtatgat	gggagggagc
	acagtgccta					
	1321	tggggagccc	aggaagctgc	tcacccaaga	tttggtgca	aaaaagtacc
35	tggagataccg					
	1381	gcaggtgccc	gacagtgatc	ccgcacgcta	tgagttcctg	tgggtccaa
	gggccttcgc					
	1441	tgaaaccagc	tatgtgaaag	tcctttagta	tgtgatcaag	gtcagtgc当地
	gagttcgctt					
40	1501	tttcttccca	tccctgcgtg	aagcagctt	gagagaggag	gaagagggag
	tctgagcatg					
	1561	agttgcagcc	aaggccagtg	ggagggggac	tggccagtg	caccttccag
	ggccgcgtcc					
45	1621	agcagcttcc	cctgcctctg	gtgacatgag	gcccattctt	cactctgaag
	agagcggtca					
	1681	gtgttctcag	tagtaggtt	ctgttctatt	gggtgacttg	gagatttatac
	tttggttctct					
	1741	tttggaaattg	ttcaaattgtt	tttttttaag	ggatgggtga	atgaacttca
	gcatccaaat					
50	1801	ttatgaatga	cagcagtcac	acagttctgt	gtatatagtt	taagggtaaag
	agtcttgcgt					
	1861	tttattcaga	ttgggaaatc	cattcttattt	tgtgaattgg	gataataaca
	gcagtggaaat					

1921 aagtacttag aaatgtaaa aatgagcagt aaaatagatg agataaagaa  
 ctaaagaaaat  
 1981 taagagatag tcaattcttgccttataacct cagtcatttc tgtaaaattt  
 ttaaagatata  
 5 2041 atgcataacct ggatttcctt ggcttctttt agaatgtaa agaaattaaa  
 tctgaataaa  
 2101 gaattcttcc tggactgg ctcttttctt ctccatgcac tgagcatctg  
 ctttttggaa  
 2161 ggcctgggt tagtagtggaa gatgctaagg taagccagac tcataacccac  
 10 ccatagggtc  
 2221 gtagagtcta ggagctgcag tcacgtaatc gaggtggcaa gatgtcctct  
 aaagatgttag  
 2281 ggaaaagtga gagaggggtg agggtgtggg gctccgggtg agagtggtgg  
 agtgtcaatg  
 15 2341 ccctgagctg gggcatttttggg aactgcagtt cttctgggg  
 gagctgattg  
 2401 taatgtatctt ggggtggatcc  
 //  
 20 L18920. Human MAGE-2 gene...[gi:436180]  
 LOCUS HUMMAGE2X 4559 bp DNA PRI  
 20-APR-1994  
 DEFINITION Human MAGE-2 gene exons 1-4, complete cds.  
 ACCESSION L18920  
 25 VERSION L18920.1 GI:436180  
 KEYWORDS  
 SOURCE Homo sapiens (human).  
 ORGANISM Homo sapiens  
 30 Eukaryota; Metazoa; Chordata; Craniata;  
 Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4559)  
 AUTHORS De Smet,C., Lurquin,C., van der Bruggen,P., De  
 35 Plaen,E.,  
 Brasseur,F. and Boon,T.  
 TITLE Sequence and expression pattern of the human MAGE2 gene  
 JOURNAL Immunogenetics 39 (2), 121-129 (1994)  
 MEDLINE 94102805  
 40 FEATURES Location/Qualifiers  
 source 1..4559  
 /organism="Homo sapiens"  
 /isolate="MZ-2"  
 45 /db\_xref="taxon:9606"  
 /map="X"  
 /haplotype="A1 A29 B37 B44 C6"  
 /sex="female"  
 /cell\_line="MZ2-MEL.43"  
 /cell\_type="melanoma cell line"  
 50 gene 561..4496  
 /gene="MAGE-2"  
 exon 561..754  
 /gene="MAGE-2"  
 /number=1  
 55 intron 755..1850  
 /gene="MAGE-2"  
 /number=1  
 exon 1851..1969

```

                /gene="MAGE-2"
                /number=2
5           intron    1970..2788
                /gene="MAGE-2"
                /number=2
               exon     2789..2854
                /gene="MAGE-2"
                /number=3
10          intron    2855..2934
                /gene="MAGE-2"
                /number=3
               exon     2935..4496
                /gene="MAGE-2"
                /number=4
15          CDS      3000..3944
                /gene="MAGE-2"
                /codon_start=1
                /protein_id="AAA17729.1"
                /db_xref="GI:436181"
20
/translation="MPLEQRSPHQCKPEEGLEARGEALGLVGAQAPATEEQQTASSSSLVEVT
LGEVPAADSPSPPHSPQGASSFSTTINYTLWRQSDEGSSNQEEGPRMFPDLE
SEFQAAISRKMVELVHFLLKYRAREPVTKAEMLESVLRNCQDFFPVIFSKASEYLQLVFGIE
VVEVVPISHLYILVTCGLSYDGLLGDNQVMPKTGLLIVLAAIAIEGDCAPEEKIWEELSM
25
EVFEGREDSVFAHPRKLLMQDLVQENYLEYRQVPGSDPACYEFLWGPRALIETSYVKVLHHTL
KIGGEPHISYPPHLHERALREGE"
polyA_signal 4479..4484
                /gene="MAGE-2"
30
BASE COUNT      1049 a    1280 c    1283 g    947 t
ORIGIN
      1 attccttcat  caaacagcca  ggagtggagga  agaggaccct  cctgagttag
      61 ccaccctcac  cacatagtgg  gaccacagaa  tccagctcag  cccctttgt
      121 acacactggc  aatgatctca  ccccgagcac  accccctcccc  ccaatgcccac
      181 ctcagagtca  gagacttgg  ctgaggggag  cagacacaat  cggcagagga
      241 ggctcagtct  ggcatccaag  tcaggacctt  gaggatgac  caaaggcccc
      301 aactcccccg  accccaccag  gatctacagc  ctcaggatcc  ccgtcccaat
      361 acaccaacac  catcttcatg  cttacccca  ccccccatac  cagatcccc
      421 atccggttcc  acccttgcgg  tgaacccagg  gaagtacgg  gcccggatgt
      481 acttgacat  tggaggtcag  aggacagcga  gattctcgcc  ctgagcaacg
      541 ggccggaggga  agcaggcgca  ggctccgtga  ggaggcaagg  taagacgcgg
      601 gaggcggggcc  tcaccccaga  cagagggccc  ccaataatcc  agcgctgcct
      661 gcctggacca  ccctgcagg  gaagacttct  caggctcagt  cgccaccacc
      721 accccccggcc  gctttaaccg  cagggaaactc  tggcgtaaga  gctttgtgtg
      781 ggctggtttag  aagtgctcag  ggcccagact  cagccaggaa  tcaaggcgtcag
      gaccccaaga

```

	841	ggggactgag	ggcaacccac	cccctaccct	caactaccaat	cccatcccc
	aacaccaacc	901	ccacccccc	ccctcaaaca	ccaacccac	ccccaaaccc
5	tcctcccca	961	ccaccatect	ggcagaatcc	ggcttgc	ctgcaatcaa
	ctccggaaat	1021	ggcgcccaag	cacgoggatc	ctgacgttca	catgtacggc
	aagggttgg	1081	gtctcgtag	tatggcctt	gggatgcaga	ggaagggccc
10	gaagacagtg	1141	gagtccctag	gggacccagc	atgccaggac	agggggccca
	gtctcaaact	1201	gagccacctt	ttcattcagc	cgagggaaatc	ctaggatgc
15	cagcaggggg	1261	ttggggccca	gcctgcgagg	agtcaagggg	aggaagaaga
	aggggacatt	1321	ggagtccaga	tcagtggcaa	ccttgggctg	gggatcctg
20	ccgaatgtgc	1381	ccctgctca	ttgcaccc	agggtgacag	ggcacagtgg
	gagggctggg	1441	acttcagg	agcagaggg	ggaatcccag	atctgccgg
25	tgcccccttc	1501	atgaggactg	gggataaccc	cggcccagaa	agaaggatg
	tggaagtccc	1561	ttgttcttag	ctctgggg	acctgatcag	gatggccct
	ctcatttga	1621	ccacaggcag	gaggttggg	aaccctcagg	aagtgacaat
30	agaggagctg	1681	tctgctcatt	tcagggg	gggggttgag	aaagggcagt
	agtaaagatg	1741	agtaacccac	aggaggccat	cataacgtt	ccctggcagg
	agccctggac	1801	aacgcacgtg	gggttaacag	gatgtggccc	atctcactt
35	atctcaggga	1861	gttgcgtacc	ttgtttcag	aaggtgactc	gttgcgtacc
	tctggc	1921	agatgcagt	gttctaggat	ctgccaagca	gttgcgtacc
	gac	1981	ggtacccctg	ggccagaatg	cagcaaggg	gttgcgtacc
40	gttagggag	2041	ttacttcaga	gaccctggc	aggcgttca	gttgcgtacc
	ctggatctt	2101	tgtatgtcagg	gaaggggagg	ccttggtctg	aaaggctgg
45	gttagggag	2161	ggtctcaggc	cctgccagga	gtggacgtga	gttgcgtacc
	ccaggacacc	2221	tggactccaa	tgaatttgg	catctctcg	gttgcgtacc
50	ggtcacgtat	2281	ggccagatgt	gggtccctc	atatccttct	gttgcgtacc
	gttcttgaca	2341	ttagatgtt	tcaagccagc	aaaagggtgg	gttgcgtacc
	aagggtgaggg	2401	ccctgagtga	gcacagaggg	gaccctccac	gttgcgtacc
55	acggaggtctg	2461	gccaacccctg	ctgagacttc	tggaaatccg	gttgcgtacc
	acactgaagg	2521	ccctgtcatt	cctctccag	gaatcaggag	gttgcgtacc
	ggccttggc					

2581 tgagtcatgt tcctcaggc acagagcaga ggggacgcag acagtgc当地  
 2641 ttgcctggaa tgcacaccaa gggcccccacc cgcccagaac aaatgggact  
 5 ccagaggggcc  
 2701 tggcctcacc ctccttattc tcagtcctgc agcctgagca tgtgctggcc  
 ggctgtaccc  
 2761 tgaggtgccc tcccacttcc tccttcaggt tctgaggggg acaggctgac  
 aagttaggacc  
 2821 cgaggcactg gaggagcatt gaaggagaag atctgttaagt aagcctttgt  
 10 cagagcctcc  
 2881 aaggttcagt tcagttctca cctaaggcct cacacacgct cttctctcc  
 ccaggcctgt  
 2941 gggtcttcat tgcccagtc ctgcccgcac tcctgcctgc tgccctgacc  
 agagtc当地  
 15 3001 tgctcttga gcagaggagt cagcactgca agcctgaaga aggccttgag  
 gccc当地  
 3061 aggccctggg cctgggggt ggc当地  
 3121 cctcttctac tctagtgaa gttaccctgg gggaggtgcc tgctgccgac  
 20 tcaccgagtc  
 3181 ctccccacag tcctcaggga gcctccagct tctcgactac catcaactac  
 actcttgg  
 3241 gacaatccga tgagggctcc agcaaccaag aagaggaggg gccaagaatg  
 tttcccgacc  
 3301 tggagtccga gttccaagca gcaatcagta ggaagatggt tgagttgg  
 25 cattttctgc  
 3361 tcctcaagta tcgagccagg gagccggtca caaaggcaga aatgctggag  
 agtgc当地  
 3421 gaaattgcca ggacttctt cccgtgatct tcagcaaagc ctccgagtac  
 30 ttgcagctgg  
 3481 tctttggcat cgaggtggtg gaagtggtcc ccatcagcca ctgtacatc  
 cttgtcacct  
 3541 gcctgggcct ctctacgat ggctgctgg ggc当地  
 3601 tcctgataat cgtcctggcc ataatcgaa tagagggcga ctgtgcccc  
 gaggagaaaa  
 3661 tctgggagga gctgagttatg ttggaggtgt ttgaggggag ggaggacagt  
 40 gtcttcgcac  
 3721 atcccaggaa gctgctcatg caagatctgg tgcaggaaaa ctacctggag  
 taccggcagg  
 3781 tgcccgccag tgatcctgca tgctacgat tccctgtgggg tccaaggggcc  
 45 ctcattgaaa  
 3841 ccagctatgt gaaagtccctg caccatacac taaagatcgg tggagaacct  
 cacatttcct  
 3901 acccaccct gcatgaacgg gctttgagag agggagaaga gtgagtctca  
 gcacatgttgc  
 3961 cagccaggcc cagtggagg gggctgggc cagtgcaccc tccagggccc  
 50 catccattag  
 4021 ctccactgc ctcgtgtat atgaggccca ttccctgcctc tttgaagaga  
 gcagtc当地  
 4081 ttcttagcag tgagttctg ttctgttggaa tgactttgag atttatctt  
 ctttcctgtt  
 4141 ggaattgttc aaatgttccct ttaacaaat ggttggatga acttcagcat  
 55 ccaaggtttat  
 4201 gaatgacagt agtcacacat agtgc当地  
 tcctgtttt  
 4261 tattcagatt gggaaatcca ttccatgggg tgagttgtca cataataaca  
 gcagtgaaat

4321 atgtatttgc ctatattgtg aacgaattag cagtaaaata catgatacaa  
 ggaactcaa  
 4381 agatagttaa ttcttgccctt atacctcagt ctattatgt aaaaaaaa  
 tatgtgtatg  
 4441 ttttgcttc tttgagaatg caaaagaaaat taaatctgaa taaattcttc  
 ctgttcaactg  
 4501 gctcatttct ttaccattca ctcagcatct gctctgtgga aggcctgg  
 agtagtggg  
 //  
 10 U03735. Human MAGE-3 anti...[gi:468825]  
 LOCUS HSU03735 4204 bp DNA PRI  
 07-APR-1994  
 DEFINITION Human MAGE-3 antigen (MAGE-3) gene, complete cds.  
 15 ACCESSION U03735  
 VERSION U03735.1 GI:468825  
 KEYWORDS .  
 SOURCE human.  
 20 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata;  
 Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini;  
 Hominidae; Homo.  
 25 REFERENCE 1 (bases 1 to 4204)  
 AUTHORS Gaugler,B., Van den Eynde,B., van der Bruggen,P.,  
 Romero,P.,  
 Gaforio,J.J., De Plaen,E., Lethe,B., Brasseur,F.  
 and Boon,T.  
 30 TITLE Human gene MAGE-3 codes for an antigen recognized on  
 a melanoma by  
 autologous cytolytic T lymphocytes  
 JOURNAL J. Exp. Med. 179, 921-930 (1994)  
 MEDLINE 94157413  
 35 REFERENCE 2 (bases 1 to 4204)  
 AUTHORS Gaugler,B.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-NOV-1993) Beatrice Gaugler, Ludwig  
 Institute for  
 40 Cancer Research, 74 Avenue Hippocrate, Brussels  
 1200, Belgium  
 FEATURES Location/Qualifiers  
 source 1..4204  
 /organism="Homo sapiens"  
 /isolate="patient M22"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /sex="female"  
 /cell\_type="lymphocyte"  
 /tissue\_type="blood"  
 /dev\_stage="adult"  
 45 50 exon 434..468  
 /number=1  
 exon 2254..2319  
 /number=2  
 55 exon 2400..3978  
 /number=3  
 gene 2465..3409  
 /gene="MAGE-3"

CDS 2465..3409  
 /gene="MAGE-3"  
 /codon\_start=1  
 /product="MAGE-3 antigen"  
 5 /protein\_id="AAA17446.1"  
 /db\_xref="GI:468826"  
 /translation="MPLEQRQSQHCKPEEGLEARGEALGLVGAQAPATEEQEAASSSTLVEVTLGE  
 VPAAESPDPPQSPQGASSLPTTMNYPLWSQSYYEDSSNQEEGPSTFPDLESEFQAALSRKVAELVH  
 10 FLLLKYRAREPTVKAEMLGSVVGWNQYFFPVIFSKASSLQLVFGIELMEVDPIGHLYIFATCLGL  
 SYDGLLGDNQIMPKAGLLIIVLAIAREGDCAPEEKIWEELSLEVPEGREDLSILGDPKLLTQHF  
 VQENYLEYRQVPGSDPACYEFLWGPRALVETSYVKVLHHMVKISGGPHISYPPLHEWLREGEE"  
 polyA\_signal 3958..3963  
 polyA\_site 3979  
 15 BASE COUNT 944 a 1144 c 1223 g 893 t  
 ORIGIN  
 1 acgcaggcag tgatgtcacc cagaccacac cccttccccc aatgccactt  
 cagggggtag  
 20 61 tcagagtcag agacttggtc tgaggggagc agaagcaatc tgcagaggat  
 ggcgggtccag  
 121 gctcagccag gcataaactt caggaccctg agggatgacc gaaggccccg  
 cccacccacc  
 181 cccaaactccc ccgacccac caggatctac agcctcagga ccccccgtccc  
 aatccttacc  
 25 241 ctttgcacca tcaccatctt catgcttacc tccaccccca tccgatcccc  
 atccaggcag  
 30 301 aatccagttc caccctgac ccgaacccag ggttagtaccg ttgccaggat  
 gtgacggcac  
 361 tgacttgcgc attggaggtc agaagaccgc gagattctcg ccctgagcaa  
 cgagcgcacgg  
 421 cctgacgtcg gcggaggggaa gccggcccaag gtcgggtgag gaggcaagg  
 aagacgctga  
 481 gggaggactg aggcggccct cacctcagac agagggcctc aaataatcca  
 gtgctgcctc  
 541 tgctgcggg cctggccac cccgcaggaa aagacttcca ggctgggtcg  
 ccactacctc  
 601 accccgcga ccccccgcgc tttagccacg gggaaactctg gggacagagc  
 ttaatgtggc  
 661 cagggcaggg ctggtagaa gaggtcaggg cccacgctgt ggcaggaatc  
 aaggtcagga  
 721 ccccgagagg gaactgaggg cagcctaacc accacccctca ccaccattcc  
 cgtcccccaa  
 781 cacccaaacc caccctatc ccccatcccc atccccaccc ccacccctat  
 cctggcagaa  
 841 tccgggcttt gcccctggta tcaagtacg gaagctccgg gaatggcggc  
 caggcacgtg  
 901 agtcctgagg ttcacatcta cggctaaggg agggaaagggg ttccgtatcg  
 cgagtatggc  
 961 cgttgggagg cagcggaaagg gcccaggccct cctggaaagac agtggagtcc  
 50 tgaggggacc  
 1021 cagcatgcca ggacaggggg cccactgtac ccctgtctca aaccgaggca  
 cctttcatt  
 1081 cggctacggg aatcctaggg atgcagaccc acttcagcag ggggttgggg  
 cccagccctg  
 1141 cgaggagtca tggggaggaa gaagagggag gactgagggg accttggagt  
 ccagatcagt  
 1201 ggcaaccttg ggctggggga tgctgggcac agtggccaaa tgtgctctgt  
 gctcattgcg

1261 ctttcagggt gaccagagag ttgagggtcg tggctgaaag agtggggactt  
 caggtcagca  
 1321 gagggaggaa tcccaggatc tgcagggccc aaggtgtacc cccaaggggc  
 ccctatgtgg  
 5 1381 tggacagatg cagtggctt aggatctgcc aagcatccag gtgaagagac  
 tgaggggagga  
 1441 ttgagggtac ccctggaca gaatgcggac tggggggccc ataaaaatct  
 gccctgtccc  
 1501 tgctgttacc tcagagagcc tgggcagggc tgtcagctga ggtccctcca  
 10 ttatccttagg  
 1561 atcactgtatc tcagggaaagg ggaagccttg gtctgagggg gctgcactca  
 gggcagtaga  
 1621 gggaggctct cagaccctac taggagtggg ggtgaggacc aagcagtctc  
 ctcacccagg  
 1681 gtacatggac ttcaataaaat ttggacatct ctcgttgc tttccggag  
 gacctgggaa  
 1741 tgtatggcca gatgtggtc ccctcatgtt tttctgtacc atatcaggtt  
 tgtgagttct  
 1801 tgacatgaga gattctcagg ccagcagaag ggagggatta ggccctataaa  
 20 ggagaaagg  
 1861 gagggccctg agtgagcaca gaggggatcc tccacccag tagagtgggg  
 acctcacaga  
 1921 gtctggccaa ccctcctgac agttctgggaa atccgtggct gcgtttgtc  
 tctgcacatt  
 1981 gggggccctg ggattcctct cccaggaatc aggagctcca ggaacaaggc  
 agtgaggact  
 2041 tggctgagg cagtgcctc aggtcacaga gtagaggggg ctcagatagt  
 gccaacggtg  
 2101 aaggtttgc accaaggggcc ccacctgccc cagaacacat  
 30 ggactccaga  
 2161 ggcgcctggcc tcaccctcaa tactttcagt cctgcagcct cagcatgcgc  
 tggccggatg  
 2221 taccctgagg tgccctctca cttccctcctt caggttctga ggggacaggc  
 tgacctggag  
 2281 gaccagaggc ccccgagga gcactgaagg agaagatctg taagtaagcc  
 tttgttagag  
 2341 ctcctaagg tccattcagt actcagctga ggtctctcac atgctccctc  
 tctccccagg  
 40 2401 ccagtgggtc tccattgccc agctcctgcc cacactcccg ctgttgccc  
 tgaccagagt  
 2461 catcatgcct cttgagcaga ggagtcaagca ctgcaaggct gaagaaggcc  
 ttgaggcccg  
 2521 aggagaggcc ctgggcctgg tgggtgcgcgaa ggctctgtc actgaggaggc  
 aggaggctgc  
 2581 ctcctcctct tctactctag ttgaagtcac cctgggggag gtgcctgtc  
 ccgagtcacc  
 2641 agatcctccc cagagtccctc agggagcctc cagoctcccc actaccatga  
 actaccctct  
 2701 ctggagccaa tcctatgagg actccagcaa ccaagaagag gagggggccaa  
 50 gcacccccc  
 2761 tgacctggag tccgagttcc aagcagcaact cagtaggaag gtggccgagt  
 tggttcattt  
 2821 tctgctcctc aagtatcgag ccagggagcc ggtcacaaaag gcagaaaatgc  
 tggggaggtgt  
 2881 cgtcggaaat tggcagttt tctttctgt gatcttcagc aaagcttcca  
 55 gttccttgca  
 2941 gctggtcttt ggcacgc gatggaaat ggacccatc ggccacttgt  
 acatctttgc

3001 caccctgcctg ggcctctcct acgatggcct gctgggtgac aatcagatca  
 tgcccaaggc  
 3061 aggccctcctg ataatcgta tggccataat cgcaagagag ggcgactgtg  
 cccctgagga  
 5 3121 gaaaatctgg gaggagctga gtgtgttaga ggtgttttag gggagggaaag  
 acagtatctt  
 3181 gggggatccc aagaagctgc tcacccaaca tttcgtgcag gaaaactacc  
 tggagtaccc  
 10 3241 gcaggtcccc ggcagtgatc ctgcattgtt tgaattcctg tggggtccaa  
 gggccctcg  
 3301 tgaaaccagc tatgtgaaag tcctgcacca tatgtaaag atcagtggag  
 gacccatcat  
 3361 ttcctaccca cccctgcatt agtgggttt gagagagggg gaagagttag  
 tctgagcacg  
 15 3421 agttgcagcc agggccagtg ggagggggtc tggggcagtg caccttccgg  
 ggcgcacatcc  
 3481 cttagtttcc actgcctcct gtgacgttag gcccattctt cactcttga  
 agcgagcagt  
 20 3541 cagcattctt agtagtgggt ttctgttctg ttggatgact ttgagattat  
 tcttttttc  
 3601 ctgttggagt tggtaaatg ttcttttaa cggatggttg aatgagcgtc  
 agcatccagg  
 3661 tttatgaatg acagtagtca cacatagtgc tggatata gtttaggagt  
 aagagtcttg  
 25 3721 ttttttactc aaattggaa atccattcca ttttgtaat tgtgacataa  
 taatacgat  
 3781 ggtaaaagta ttgtttaaa attgtgagcg aattagcaat aacatacatg  
 agataactca  
 3841 agaaatcaaa agatagttga ttcttgcctt gtacctaatttctgt  
 aaattaaaca  
 3901 aatatgcaaa ccaggatttc cttgacttct ttgagaatgc aagcgaaatt  
 aaatctgaat  
 3961 aaataattct tcctttcac tggctcgttt cttttccgtt cactcagcat  
 ctgctctgtg  
 40 4021 ggaggccctg ggttagtagt ggggatgcta aggttaagcca gactcacgcc  
 taccatagg  
 4081 gctgttagagc ctaggacctg cagtcataata attaaggtagg tgagaagtcc  
 tgtaagatgt  
 4141 agagggaaatg taagagaggg gtgagggtgt ggcgctccgg gtgagagtag  
 tggagtgtca  
 4201 gtgc  
 //

**AF043498. Homo sapiens pros...[gi:2909843]**  
 45 LOCUS AF043498 990 bp mRNA PRI 24-  
 FEB-1998  
 DEFINITION Homo sapiens prostate stem cell antigen (PSCA)  
 mRNA, complete cds.  
 ACCESSION AF043498  
 50 VERSION AF043498.1 GI:2909843  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 55 Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 Homo.  
 REFERENCE 1 (bases 1 to 990)

AUTHORS Reiter,R.E., Gu,Z., Watabe,T., Thomas,G., Kinga,S.,  
 Davis,E.,  
 Wahl,M., Nisitani,S., Yamashiro,J., Le Beau,M.M.,  
 Losa,M. and  
 5 Witte,O.N.  
 TITLE Prostate stem cell antigen: a cell surface marker  
 overexpressed in  
 prostate cancer  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (4), 1735-1740 (1998)  
 10 MEDLINE 98132661  
 REFERENCE 2 (bases 1 to 990)  
 AUTHORS Reiter,R.E.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JAN-1998) Urology, UCLA, 66-134 CHS  
 15 10833 Le Conte  
 Ave., Los Angeles, CA 90095, USA  
 FEATURES Location/Qualifiers  
 source 1..990  
 /organism="Homo sapiens"  
 20 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8q24.2"  
 /note="LAPC-4 prostate cancer xenograft"  
 gene 1..990  
 25 /gene="PSCA"  
 CDS 18..389  
 /gene="PSCA"  
 /note="GPI-anchored cell surface protein"  
 /codon\_start=1  
 30 /product="prostate stem cell antigen"  
 /protein\_id="AAC39607.1"  
 /db\_xref="GI:2909844"  
  
 35 /translation="MKAVLLALLMAGLALQPQGTALLCYSCAKQVSNEDCLQVENCTQLGEQCWTARIRAVGLLTIVISKGCSLNCSVDDSQDYYVGKKNITCCDTDLCNASGAHALQPAAAILALLPALGLLIWGPQL"  
 BASE COUNT 193 a 299 c 285 g 202 t 11 others  
 ORIGIN  
 40 1 agggagagggc agtgaccatg aaggctgtgc tgcttgccct gttgatggca  
 ggcttggccc  
 61 tgcagccagg cactgccctg ctgtgtact cctgcaaagc ccaggtgagc  
 aacgaggact  
 121 gcctgcaggt ggagaactgc acccagctgg gggagcagtg ctggaccgcg  
 cgcattccgcg  
 45 181 cagttggcct cctgaccgtc atcagcaaag gctgcagctt gaactgcgtg  
 gatgactcac  
 241 aggactacta cgtggcaag aagaacatca cgtgctgtga caccgacttg  
 tgcaacgcca  
 301 gcggggccca tgccctgcag cccgctgccg ccattcctgc gctgctccct  
 50 gcactcggcc  
 361 tgctgctctg gggacccggc cagctataagg ctctgggggg ccccgctgca  
 gccccacactg  
 421 ggtgtgggtgc cccaggcctt tgtgccactc ctcacagaac ctggcccaagt  
 gggagcctgt  
 481 cctggttcct gaggcacatc ctaacgcaag tttgaccatg tatgtttgca  
 cccctttcc  
 55 541 ccnaaccctg accttccat gggccttttc caggattccn accnggcaga  
 tcagtttag

601 tganacanat ccgcntgcag atggcccctc caaccnnttn tggtgntgtt  
 tccatggccc  
 661 agcattttcc acccttaacc ctgtgttcag gcacttnttc ccccaggaag  
 ccttccctgc  
 721 ccacccatt tatgaattga gccagggttg gtccgtggtg tcccccgac  
 ccagcagggg  
 781 acaggcaatc aggagggccc agtaaaggct gagatgaagt ggactgagta  
 gaactggagg  
 841 acaaagagttg acgtgagttc ctgggagttt ccagagatgg ggcctggagg  
 10 cctggaggaa  
 901 ggggccaggg ctcacatttg tgggntccc gaatggcagc ctgagcacag  
 cgtaggccct  
 961 taataaacac ctgttggata agccaaaaaa  
 //  
 15

**P06870. GLANDULAR KALLIKR...[gi:125170]**  
 LOCUS KLK1\_HUMAN 262 aa PRI 20-  
 AUG-2001  
 DEFINITION GLANDULAR KALLIKREIN 1 PRECURSOR (TISSUE  
 20 KALLIKREIN)  
 (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).  
 ACCESSION P06870  
 PID g125170  
 VERSION P06870 GI:125170  
 25 DBSOURCE swissprot: locus KLK1\_HUMAN, accession P06870;  
 class: standard.  
 extra accessions: Q9UMJ1, created: Jan 1, 1988.  
 sequence updated: Jan 1, 1988.  
 annotation updated: Aug 20, 2001.  
 30 xrefs: gi: gi: 186652, gi: gi: 186653, gi: gi: 186649,  
 gi: gi:  
 186651, gi: gi: 186645, gi: gi: 186646, gi: gi:  
 186647, gi: gi:  
 186648, gi: gi: 34026, gi: gi: 34027, gi: gi: 186643,  
 35 gi: gi:  
 386843, gi: gi: 67558  
 xrefs (non-sequence databases): HSSP P00757, MEROPS  
 S01.160,  
 40 GlycoSuiteDB P06870, MIM 147910, InterPro IPR001314,  
 InterPro  
 IPR001254, Pfam PF00089, PRINTS PR00722, PROSITE  
 PS50240, PROSITE  
 PS00134, PROSITE PS00135  
 45 KEYWORDS Hydrolase; Serine protease; Glycoprotein; Multigene  
 family;  
 Zymogen; Signal.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 50 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 HOMO.  
 REFERENCE 1 (residues 1 to 262)  
 AUTHORS Fukushima,D., Kitamura,N. and Nakanishi,S.  
 55 TITLE Nucleotide sequence of cloned cDNA for human  
 pancreatic kallikrein  
 JOURNAL Biochemistry 24, 8037-8043 (1985)  
 REMARK SEQUENCE FROM N.A.

5 TISSUE=Pancreas  
 REFERENCE 2 (residues 1 to 262)  
 AUTHORS Evans,B.A., Yun,Z.X., Close,J.A., Tregear,G.W.,  
 Kitamura,N., Nakanishi,S., Callen,D.F., Baker,E., Hyland,V.J.,  
 Sutherland,G.R. and Richards,R.I.  
 10 TITLE Structure and chromosomal localization of the human  
 renal kallikrein gene  
 JOURNAL Biochemistry. 27 (9), 3124-3129 (1988)  
 MEDLINE 88269498  
 PUBMED 2898948  
 15 REMARK SEQUENCE FROM N.A.  
 TISSUE=Kidney  
 REFERENCE 3 (residues 1 to 262)  
 AUTHORS Angermann,A., Bergmann,C. and Appelhans,H.  
 TITLE Cloning and expression of human salivary-gland  
 20 kallikrein in Escherichia coli  
 JOURNAL The Biochemical journal. 262 (3), 787-793 (1989)  
 MEDLINE 90073574  
 PUBMED 2686621  
 25 REMARK SEQUENCE FROM N.A.  
 TISSUE=Salivary gland  
 REFERENCE 4 (residues 1 to 262)  
 AUTHORS Baker,A.R. and Shine,J.  
 TITLE Human kidney kallikrein: cDNA cloning and sequence  
 30 analysis  
 JOURNAL DNA (Mary Ann Liebert, Inc.) 4 (6), 445-450 (1985)  
 MEDLINE 86135264  
 PUBMED 3853975  
 REMARK SEQUENCE OF 17-262 FROM N.A.  
 35 TISSUE=Kidney  
 REFERENCE 5 (residues 1 to 262)  
 AUTHORS Lu,H.S., Lin,F.K., Chao,L. and Chao,J.  
 TITLE Human urinary kallikrein. Complete amino acid sequence  
 40 and sites of glycosylation  
 JOURNAL International journal of peptide and protein research.  
 33 (4), 237-249 (1989)  
 MEDLINE 89326688  
 PUBMED 2666327  
 45 REMARK SEQUENCE OF 25-262.  
 TISSUE=Urine  
 REFERENCE 6 (residues 1 to 262)  
 AUTHORS Kellermann,J., Lottspeich,F., Geiger,R. and  
 50 Deutzmann,R.  
 TITLE Human urinary kallikrein--amino acid sequence and  
 carbohydrate attachment sites  
 JOURNAL Protein sequences & data analysis. 1 (3), 177-182  
 (1988)  
 MEDLINE 88203586  
 PUBMED 3163150  
 55 REMARK SEQUENCE OF 25-262, AND CARBOHYDRATE-LINKAGE SITES.  
 TISSUE=Urine

REFERENCE 7 (residues 1 to 262)  
 AUTHORS Lottspeich,F., Geiger,R., Henschen,A. and Kutzbach,C.  
 TITLE N-Terminal amino acid sequence of human urinary  
 kallikrein homology  
 5 with other serine proteases  
 JOURNAL Hoppe-Seyler's Zeitschrift fur physiologische Chemie.  
 360 (12), 1947-1950 (1979)  
 MEDLINE 80114126  
 10 PUBMED 393608  
 REMARK SEQUENCE OF 25-55.  
 TISSUE=Urine  
 REFERENCE 8 (residues 1 to 262)  
 AUTHORS Takahashi,S., Irie,A., Katayama,Y., Ito,K. and  
 15 Miyake,Y.  
 TITLE N-terminal amino acid sequence of human urinary  
 prokallikrein  
 JOURNAL Journal of biochemistry. 99 (3), 989-992 (1986)  
 MEDLINE 86223893  
 20 PUBMED 3635530  
 REMARK SEQUENCE OF 28-47.  
 TISSUE=Urine  
 [FUNCTION] GLANDULAR KALLIKREINS CLEAVE MET-LYS AND  
 ARG-SER BONDS  
 25 IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 [CATALYTIC ACTIVITY] PREFERENTIAL CLEAVAGE OF ARG-|-  
 XAA BONDS IN  
 SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO  
 RELEASE  
 30 KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES  
 HYDROLYSIS OF  
 MET-|-XAA OR LEU-|-XAA.  
 [SIMILARITY] BELONGS TO PEPTIDASE FAMILY S1; ALSO  
 KNOWN AS THE  
 35 TRYPSIN FAMILY. KALLIKREIN SUBFAMILY.  
 FEATURES Location/Qualifiers  
 source 1..262  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 40 Protein 1..262  
 /product="GLANDULAR KALLIKREIN 1 PRECURSOR"  
 /EC\_number="3.4.21.35"  
 Region 1..18  
 /region\_name="Signal"  
 /note="PROBABLE."  
 45 Region 19..24  
 /region\_name="Propeptide"  
 /note="ACTIVATION PEPTIDE (PROBABLE)."  
 50 Region 25..262  
 /region\_name="Mature chain"  
 /note="GLANDULAR KALLIKREIN 1."  
 Bond bond(31,174)  
 /bond\_type="disulfide"  
 /note="BY SIMILARITY."  
 55 Bond bond(50,66)  
 /bond\_type="disulfide"  
 /note="BY SIMILARITY."

```

      Site      65
      Site      /site_type="active"
      Site      /note="CHARGE RELAY SYSTEM."
      Site      93
      Site      /site_type="glycosylation"
      Site      /note="O-LINKED."
      Site      102
      Site      /site_type="glycosylation"
      Site      /note="N-LINKED (GLCNAC...)."
      Site      104
      Site      /site_type="glycosylation"
      Site      /note="O-LINKED."
      Site      108
      Site      /site_type="glycosylation"
      Site      /note="N-LINKED (GLCNAC...)."
      Site      120
      Site      /site_type="active"
      Site      /note="CHARGE RELAY SYSTEM."
      Region    145
      Region    /region_name="Variant"
      Bond      /note="Q -> E. /FTId=VAR_006625."
      Bond      bond(153,220)
      Bond      /bond_type="disulfide"
      Bond      /note="BY SIMILARITY."
      Site      165
      Site      /site_type="glycosylation"
      Site      /note="N-LINKED (GLCNAC...) (PARTIAL)."
      Site      167
      Site      /site_type="glycosylation"
      Site      /note="O-LINKED."
      Bond      bond(185,199)
      Bond      /bond_type="disulfide"
      Bond      /note="BY SIMILARITY."
      Region    186
      Region    /region_name="Variant"
      Bond      /note="E -> K. /FTId=VAR_006626."
      Bond      bond(210,235)
      Bond      /bond_type="disulfide"
      Bond      /note="BY SIMILARITY."
      Site      214
      Site      /site_type="active"
      Site      /note="CHARGE RELAY SYSTEM."
      ORIGIN
      1 mwflvlclal slggtaapp iqsrivggwe ceqhsqpwqa alyhfstfqc
      45 ggilvhraqwv
      61 ltaahcisdn yqlwlgrhnl fddentaqfv hvsesfphpg fnmsllenht
      rqadedyshd
      121 lmlrltpea dtitdavkvv elptqepevg stclasgwgs iepenfsfpd
      dlqcvdlkil
      50 181 pndecekahv qkvtdfmlcv ghleggkdtc vgdsggplmc dgvlqgvtsw
      gyvpcgtpnk
      241 psvavrvlsy vkwiedtiae ns
      //
      55 P08217. ELASTASE 2A PRECU...[gi:119255]
      LOCUS      EL2A_HUMAN    269 aa
      AUG-2001

```

DEFINITION ELASTASE 2A PRECURSOR.  
 ACCESSION P08217  
 PID g119255  
 VERSION P08217 GI:119255  
 5 DBSOURCE swissprot: locus EL2A\_HUMAN, accession P08217;  
     class: standard.  
     created: Aug 1, 1988.  
     sequence updated: Aug 1, 1988.  
     annotation updated: Aug 20, 2001.  
 10 xrefs: gi: gi: 182022, gi: gi: 182023, gi: gi: 182057,  
     gi: gi:  
     182058, gi: gi: 88298, gi: gi: 88299  
     xrefs (non-sequence databases): MEROPS S01.155,  
 15 InterPro IPR001314,  
     InterPro IPR001254, Pfam PF00089, PRINTS PR00722,  
 PROSITE PS50240,  
     PROSITE PS00134, PROSITE PS00135  
 KEYWORDS Hydrolase; Serine protease; Pancreas; Zymogen; Signal.  
 SOURCE human.  
 20 ORGANISM Homo sapiens  
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
     Euteleostomi;  
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
     Homo.  
 25 REFERENCE 1 (residues 1 to 269)  
     AUTHORS Kawashima,I., Tani,T., Shimoda,K. and Takiguchi,Y.  
     TITLE Characterization of pancreatic elastase II cDNAs: two  
     elastase II  
 30 JOURNAL mRNAs are expressed in human pancreas  
     DNA 6 (2), 163-172 (1987)  
     MEDLINE 87217962  
     REMARK SEQUENCE FROM N.A.  
 REFERENCE 2 (residues 1 to 269)  
 35 AUTHORS Fletcher,T.S., Shen,W.F. and Largman,C.  
     TITLE Primary structure of human pancreatic elastase 2  
     determined by  
     sequence analysis of the cloned mRNA  
     JOURNAL Biochemistry 26 (23), 7256-7261 (1987)  
     MEDLINE 88107669  
 40 REMARK SEQUENCE FROM N.A.  
     [FUNCTION] ACTS UPON ELASTIN.  
     [CATALYTIC ACTIVITY] PREFERENTIAL CLEAVAGE: LEU-| -XAA,  
     MET-| -XAA  
     AND PHE-| -XAA. HYDROLYSES ELASTIN.  
 45 [SUBCELLULAR LOCATION] SECRETED.  
     [TISSUE SPECIFICITY] PANCREAS.  
     [SIMILARITY] BELONGS TO PEPTIDASE FAMILY S1; ALSO  
     KNOWN AS THE  
     TRYPSIN FAMILY. ELASTASE SUBFAMILY.  
 50 FEATURES Location/Qualifiers  
     source 1..269  
           /organism="Homo sapiens"  
           /db\_xref="taxon:9606"  
     Region 1..16  
           /region\_name="Signal"  
     Protein 1..269  
           /product="ELASTASE 2A PRECURSOR"  
           /EC\_number="3.4.21.71"

```

      Region      17..28
      /region_name="Propeptide"
      /note="ACTIVATION PEPTIDE."
5       Region      29..269
      /region_name="Mature chain"
      /note="ELASTASE 2A."
      Bond      bond(58,74)
      /bond_type="disulfide"
      /note="BY SIMILARITY."
10      Site      73
      /site_type="active"
      /note="CHARGE RELAY SYSTEM (BY SIMILARITY)."
      Site      121
      /site_type="active"
15      Bond      bond(155,222)
      /bond_type="disulfide"
      /note="BY SIMILARITY."
      Bond      bond(186,202)
20      /bond_type="disulfide"
      /note="BY SIMILARITY."
      Bond      bond(212,243)
      /bond_type="disulfide"
      /note="BY SIMILARITY."
25      Site      216
      /site_type="active"
      /note="CHARGE RELAY SYSTEM (BY SIMILARITY)."
      ORIGIN
      1 mirtlllstl vagalscgdp typpyvtrvv ggeearpnsw pwqvsdqyss
30      ngkwyhtcg
      61 slianswlt aahcisssrt yrvglgrhnl yvaesgslav svskivvhkd
      wnsnqiskgn
      121 diallkklap vsldkiqla clppagtilp nnypcyvtgw grlqtnqavp
      dvlqqgrllv
35      181 vdyatcsssa wwgssvktsm icaggdgvvis scngdsggpl ncqasdgrwq
      vhgivsfgrs
      241 lgcnyyhkps vftrvsnyid winsviann
      //

40      NP_056933. pancreatic elastase...[gi:7705648]
LOCUS      NP_056933      269 aa          PRI      02-
NOV-2000
DEFINITION  pancreatic elastase IIB [Homo sapiens].
45      ACCESSION  NP_056933
      PID        g7705648
      VERSION    NP_056933.1 GI:7705648
      DBSOURCE   REFSEQ: accession NM_015849.1
      KEYWORDS
      SOURCE     human.
50      ORGANISM  Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
                  Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
                  Homo.
55      REFERENCE 1 (residues 1 to 269)
      AUTHORS    Kawashima,I., Tani,T., Shimoda,K. and Takiguchi,Y.
      TITLE      Characterization of pancreatic elastase II cDNAs: two
                  elastase II

```

mRNAs are expressed in human pancreas

JOURNAL DNA 6 (2), 163-172 (1987)  
 MEDLINE 87217962

COMMENT PROVISIONAL REFSEQ: This record has not yet been  
 5 subject to final NCBI review. The reference sequence was derived from  
 M16653.1.

FEATURES Location/Qualifiers  
 10 source 1..269  
     /organism="Homo sapiens"  
     /db\_xref="taxon:9606"  
     /chromosome="12"  
     /map="12"  
 15 Protein 1..269  
     /product="pancreatic elastase IIB"  
     sig\_peptide 1..16  
     /note="pancreatic elastase IIB signal  
     peptide"  
 20 Region 28..262  
     /region\_name="Trypsin-like serine protease"  
     /db\_xref="CDD:Tryp\_SPC"  
     /note="Tryp\_SPC"  
     mat\_peptide 29..269  
     /product="pancreatic elastase IIB mature  
 25 peptide"  
     Region 31..262  
     /region\_name="Trypsin"  
     /db\_xref="CDD:pfam00089"  
     /note="trypsin"  
 30 CDS 1..269  
     /gene="LOC51032"  
     /db\_xref="LocusID:51032"  
     /coded\_by="NM\_015849.1:26..835"

ORIGIN

35 1 mirtlllsl vagalscgvs tyapdmsrml ggearpnsw pwqvs1qyss  
 ngqwyhtcgg  
 61 slianswvl aahcisssri yrvmllgqhn1 yvaesgslav svskivvhkd  
 wnsnqvskgm  
 121 diallklamp vsltdkiqla clppagttilp nnypcyvtgw grlqtnqalp  
 40 ddlkqgrllv  
 181 vdyatcsssg wwgstvktm icaggdgVIC tcngdsggpl ncqasdgrwe  
 vhgigsitsv  
 241 lgcnyyykps iftrvsnynd winsviann  
 //

45

PRAME

50 LOCUS NM\_006115 2148 bp mRNA PRI 19-JUN-  
 2001

DEFINITION Homo sapiens preferentially expressed antigen in  
 melanoma (PRAME),  
 mRNA.

55 ACCESSION NM\_006115  
 VERSION NM\_006115.1 GI:5174640  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 HOMO.  
 5 REFERENCE 1 (bases 1 to 2148)  
 AUTHORS Ikeda,H., Lethe,B., Lehmann,F., Van Baren,N.,  
 Baurain,J.-F., De  
 Smet,C., Chambost,H., Vitale,M., Moretta,A., Boon,T. and  
 Coulie,P.G.  
 10 TITLE Characterization of an antigen that is recognized on a  
 melanoma showing partial HLA loss by CTL expressing an NK inhibitor  
 receptor  
 JOURNAL Immunity 6 (2), 199-208 (1997)  
 MEDLINE 97199265  
 15 REFERENCE 2 (bases 1 to 2148)  
 AUTHORS Williams JM, Chen GC, Zhu L and Rest RF.  
 TITLE Using the yeast two-hybrid system to identify human  
 epithelial cell  
 20 gonococci proteins that bind gonococcal Opa proteins: intracellular  
 bind pyruvate kinase via their Opa proteins and  
 require host  
 pyruvate for growth  
 25 JOURNAL Mol. Microbiol. 27 (1), 171-186 (1998)  
 MEDLINE 98125741  
 PUBMED 9466265  
 REFERENCE 3 (bases 1 to 2148)  
 AUTHORS van Baren,N., Chambost,H., Ferrant,A., Michaux,L., Ikeda,H  
 Millard,I., Olive,D., Boon,T. and Coulie,P.G.  
 30 TITLE PRAME, a gene encoding an antigen recognized on a  
 human melanoma by  
 cytolytic T cells, is expressed in acute leukaemia cells  
 JOURNAL Br. J. Haematol. 102 (5), 1376-1379 (1998)  
 MEDLINE 98423996  
 PUBMED 9753074  
 REFERENCE 4 (bases 1 to 2148)  
 AUTHORS Dunham I, Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JI  
 Bruskiewich R, Beare DM, Clamp M, Smink LJ, Ainscough  
 40 R, Almeida JP, Babbage A, Bagguley C, Bailey J, Barlow K, Bates  
 KN, Beasley O,  
 Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J,  
 Burrill WD,  
 45 TITLE O'Brien KP and et al.  
 JOURNAL The DNA sequence of human chromosome 22  
 Nature 402 (6761), 489-495 (1999)  
 MEDLINE 20057165  
 PUBMED 10591208  
 50 REMARK Erratum: [{published erratum appears in Nature 2000 Apr  
 20;404(6780):904}].  
 COMMENT REVIEWED REFSEQ: This record has been curated by NCBI  
 staff. The  
 reference sequence was derived from U65011.1.  
 55 amino acid Summary: The protein encoded by this gene has a 509  
 antigen, lacking a signal sequence, and recognized on  
 a human

melanoma cell line by a T-lymphocyte clone. A  
 significant level of  
 this mRNA is detected in normal testis as well as in  
 many  
 5 melanomas, non-small cell lung carcinomas, sarcomas,  
 head and neck  
 tumors and renal carcinomas. The encoded protein is  
 expressed  
 predominantly in acute leukemias carrying chromosomal  
 10 abnormalities  
 such as translocation t(8:21), which fuses the AML1  
 and ETO genes.  
 Its expression shares several characteristics with the  
 expression  
 15 patterns of MAGE, BAGE, and GAGE gene families, all of  
 which are  
 expressed in tumors. This protein is expressed in a  
 higher  
 20 proportion of samples than genes of the MAGE, BAGE,  
 and GAGE  
 families, and of these four groups, only this protein  
 is expressed  
 by acute myeloid leukemias.  
 25 FEATURES Location/Qualifiers  
 source 1..2148  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 30 /map="22q11.22"  
 /haplotype="A24 A28 B13 B44 Cw6 Cw7"  
 gene 1..2148  
 /gene="PRAME"  
 /note="MAPE; OIP4"  
 35 /db\_xref="LocusID:23532"  
 /db\_xref="MIM:606021"  
 CDS 236..1765  
 /gene="PRAME"  
 /note="melanoma antigen preferentially  
 40 expressed in tumors; Opa-interacting protein OIP4"  
 /codon\_start=1  
 /db\_xref="LocusID:23532"  
 /db\_xref="MIM:606021"  
 45 /product="preferentially expressed antigen of  
 melanoma"  
 /protein\_id="NP\_006106.1"  
 /db\_xref="GI:5174641"  
 50 /translation="MERRRLWGSIQSRYISMSVWTSPRRLVELAGQSLKDEALAIAALELLPREL  
 FPPLFMAAFDGRHSQTLKAMVQAWPFTCLPLGVLMKGQHLHLETFKAVLDGLDVLLAQEVPRRRWK  
 LQVLDLRKNSHQDFWTVWSGNRASLYSFPEPEAAOPMTKRKVDGLSTEAEQPFIPVEVLVDLFLK  
 EGACDELFSYLIEKVKRKKNVLRLCCKKLKIFAMPQDIKMILKMVQLDSIEDLEVCTWKLPTLA  
 KFSPYLGQMINLRRLLSHIHASSYISPEKEEQYIAQFTSQFLSLQCLQALYVDSLFFLRGRLDQL  
 55 LRHVMNPLETLSITNCRILSEGDMHLSQSPSVSQLSVLSQLGVM LTDVSPPEPLQALLERASATLQD  
 LVFDECGITDDQLLALLPSSLHCSQLTTSFYGNNSISISALQSLQHLIGLSNLTHVLYPVPLESY  
 EDIHGTLHLERLAYLHARLRELLCELGRPSMVWL SANPCPHCGDRTFYDPEPILCPFCMPN"  
 variation 254

```

      /allele="C"
      /allele="T"
      /db_xref="dbSNP:1129172"
5       variation      698
      /allele="C"
      /allele="A"
      /db_xref="dbSNP:1063107"
10      variation      1001
      /allele="C"
      /allele="T"
      /db_xref="dbSNP:2229695"
15      variation      1498
      /allele="T"
      /allele="C"
      /db_xref="dbSNP:2229696"
20      variation      1762
      /allele="C"
      /allele="T"
      /db_xref="dbSNP:13604"
25      variation      1952
      /allele="A"
      /allele="G"
      /db_xref="dbSNP:7104"
      variation      1984
      /allele="A"
      /allele="C"
      /db_xref="dbSNP:8405"
30      variation      2043
      /allele="T"
      /allele="A"
      /db_xref="dbSNP:1804685"
      polyA_signal  2108..2113
      polyA_site    2130
      BASE COUNT    534 a    548 c    558 g    508 t
35      ORIGIN
      1 gttcagggt acagctcccc cgcagccaga agccgggcct gcagccccctc
      agcaccgctc
      61 cgggacaccc caccgcgttc ccaggcgtga cctgtcaaca gcaacttcgc
      ggtgtggta
      121 actctctgag gaaaaaccat tttgattatt actctcagac gtgcgtggca
40      acaagtgact
      181 gagacctaga aatccaagcg ttggaggtcc tgaggccagc ctaagtcgct
      tcaaaatgga
      241 acgaaggcgt ttgtgggtt ccattcagag ccgatacatc agcatgagtg
45      tgtggacaag
      301 cccacggaga cttgtggagc tggcagggca gagcctgctg aaggatgagg
      ccctggccat
      361 tgccgcctg gagttgtgc ccagggagct cttcccgcca ctcttcatgg
      cagccttta
      421 cgggagacac agccagaccc tgaaggcaat ggtgcaggcc tggcccttca
50      cctgcctccc
      481 tctgggagtg ctgatgaagg gacaacatct tcacctggag accttcaaag
      ctgtgcttga
      541 tggacttgat gtgctccttg cccaggaggt tcgccccagg aggtggaaac
55      ttcaagtgtct
      601 ggatttacgg aagaactctc atcaggactt ctggactgta tggtctggaa
      acagggccag

```

661 tctgtactca tttccagagc cagaaggcgc tcagccatg acaaagaagc  
 gaaaagtaga  
 721 tggtttgagc acagagggcag agcagccctt cattccagta gaggtgctcg  
 tagacctgtt  
 5 781 cctcaaggaa ggtgcctgtg atgaattgtt ctcctacctc attgagaaag  
 tgaagcgaaa  
 841 gaaaaatgtt ctacgcctgt gctgtaaagaa gctgaagatt tttgcaatgc  
 ccatgcagga  
 901 tatcaagatg atcctgaaaa tggtgcagct ggactctatt gaagatttgg  
 10 aagtgacttg  
 961 tacctgaaag ctacccacct tggcggaaatt ttctccttac ctggccaga  
 tgattaatct  
 1021 gcgttagactc ctcctctccc acatccatgc atcttcctac atttccccgg  
 agaaggaaga  
 15 1081 gcgttatatc gcccagttca cctctcagtt cctcagtcg cagtgcctgc  
 aggctctcta  
 1141 tgtggactct ttattttcc ttagaggccg cctggatcag ttgctcaggg  
 acgtgatgaa  
 1201 ccccttggaa accctctcaa taactaactg ccggcttgc gaaggggatg  
 20 tgatgcatt  
 1261 gtcccagagt cccagegtca gtcagctaag tgtcctgagt ctaagtgggg  
 tcatgctgac  
 1321 cgtatgtaagt cccgagcccc tccaagctct gctggagaga gcctctgcca  
 ccctccagga  
 25 1381 cctggtcttt gatgagtggtg ggatcacgga tgatcagctc cttgcctcc  
 tgccctccct  
 1441 gagccactgc tcccagctta caaccttaag cttctacggg aattccatct  
 ccatatctgc  
 1501 cttgcagagt ctcctgcagc acctcatcg gctgagcaat ctgacccacg  
 30 tgctgtatcc  
 1561 tgtccccctg gagagttatg aggacatcca tggtaccctc cacctggaga  
 ggcttgccta  
 1621 tctgcattgcc aggctcaggg agttgctgtg tgagttgggg cggcccgca  
 tggtctggct  
 35 1681 tagtgcacaaac cccgtcctc actgtggggc cagaacccatc tatgacccgg  
 agcccatcct  
 1741 gtgccccctgt ttcattgccta actagctggg tgcacatatc aaatgcttca  
 ttctgcatac  
 1801 ttggacacta aagccaggat gtgcattgcattt cttgaagcaa caaaggagcc  
 40 acagtttcag  
 1861 acaaatgttc agtgtgagtg aggaaaaacat gttcagtgag gaaaaaaacat  
 tcagacaaat  
 1921 gttcagtgag gaaaaaaagg ggaagttggg gataggcaga tggacttg  
 aggagttaat  
 45 1981 gtgatctttg gggagataca tcttatagag ttagaaatag aatctgaatt  
 tctaaaggga  
 2041 gattctggct tggaaagtac atgttaggat taatccctgt gtagactgtt  
 gtaaagaaac  
 2101 tggaaaaat aaagagaagc aatgtgaago aaaaaaaaaaa aaaaaaaaa

## ED-B domain of Fibronectin

LOCUS HSFIBEDB 2823 bp DNA linear  
 PRI 09-AUG-1999  
 5 DEFINITION Human fibronectin gene ED-B region.  
 ACCESSION X07717  
 VERSION X07717.1 GI:31406  
 KEYWORDS alternate splicing; fibronectin.  
 SOURCE human.  
 10 ORGANISM Homo sapiens  
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
     Euteleostomi;  
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
     Homo.  
 15 REFERENCE 1 (bases 1 to 2823)  
     AUTHORS Paolella,G., Henchcliffe,C., Sebastio,G. and  
     Baralle,F.E.  
     TITLE Sequence analysis and *in vivo* expression show that  
     alternative  
 20 splicing of ED-B and ED-A regions of the human  
     fibronectin gene are  
     independent events  
     JOURNAL Nucleic Acids Res. 16 (8), 3545-3557 (1988)  
     MEDLINE 88233940  
 25 FEATURES Location/Qualifiers  
     source 1..2823  
     /organism="Homo sapiens"  
     /db\_xref="taxon:9606"  
     /clone="MA10"  
 30 exon 1..104  
     /number=1  
     /product="fibronectin"  
     CDS join(<2..104,1375..1647,2758..>2823)  
     /codon\_start=1  
 35 /product="fibronectin"  
     /protein\_id="CAB52437.1"  
     /db\_xref="GI:5725425"  
  
     /translation="CTFDNLSPGLEYNVSVYTVKDDKESVPISDTIIPPEVPQLTDLSF"

VDITDSSIGLRWTPLNSSTIIGYRITVVAAGEGIPIFEDFVDSSVGYYTGTLEPGID

YDISVITLINGGESAPTTLTQQTAVPPPPTDLRFTNIGPDTMRVTW"

5           intron           105..1374  
               /number=1  
           exon           1375..1647  
               /note="ED-B exon"  
               /number=2  
 10           /product="fibronectin"  
           intron           1648..2757  
               /number=2  
           exon           2758..2823  
               /number=3  
 15           /product="fibronectin"  
 BASE COUNT    824 a    556 c    528 g    915 t  
 ORIGIN  
       1 ctgcactttt gataacctga gtcccgccct ggagtacaat gtcagtgttt  
       acactgtcaa  
 20        61 ggatgacaag gaaagtgtcc ctatctctga taccatcatc ccaggtaata  
       gaaaataaagc  
       121 tgctatcctg agagtacat tccaataaga gtggggattt gcatcttaat  
       ccccagatgc  
       181 ttaagggtgt caactatatt tgggattttt ttccgatctc ccagctgcac  
 25        tttccaaaac  
       241 caagaagtca aagcagcgtt ttggacaaaa tgcttgctgt taacactgct  
       ttactgtctg  
       301 tgcttcactg ggatgctgtg tggcagcg agtatgtttt ggagtggcag  
       ccatggcttt  
       361 aactctgtat tggcgttca catggaaatgt tgactaaaac actgtcacgt  
       gtctgtactc  
       421 agtactgata ggctcaaaatgtt aatatggtaa atgcattttt tcagttttt  
       tctgccccat  
       481 ttacaatcc atatcaatcc ccaacagctg cctatttcat cttgcagttt  
 35        caaatcccttc  
       541 tttttgaaaa ttggattttt aaaaaaaatgtt aagttttt aacacccatca  
       gggttttttct  
       601 ttcttggtggc cttgaaagac aacattgcaaa aggcctgtcc taaggatagg  
       cttggggggc

661 cattgggtta taacataatg aaagcattgg acagatcgtg tcccccttg  
gactcttcag  
721 tagaatgctt ttactaacgc taattacatg ttttgcattat gaatgaacct  
aaaatagtgg  
5 781 caatggcctt aacctaggcc tgtcttcct cagcctgaat gtgcctttga  
atggcacatt  
841 tcacaccata cattcataat gcattagcgt tatggccatg atgttgtcat  
gagttttgtt  
901 tgggagaaaaaaa aaaaatcaatt tatcacccat ttattatttt ttccgggttgt  
10 tcatgcaagc  
961 ttatttcta ctaaaacagt ttggaaatta ttaaaagcat tgctgatact  
tacttcagat  
1021 attatgtcta ggctctaaga atggttcga catcctaaac agccatatga  
tttttaggaa  
15 1081 tctgaacagt tcaaattgtt cccttaagg atgtttcaa aatgtaaaaaa  
atatatatat  
1141 atatatatat tccctaaaag aatattcctg tttattctc tagggaaagca  
aactgttcat  
1201 gatgcttagg aagtctttc agagaattta aaacagattt catattacca  
20 tcattgcttt  
1261 aacattccac caattttact actagtaacc tgatatacac tgctttatTT  
tttcctctt  
1321 tttccctctt atttccctt tgccctcccccc tccctttgct ttgttaactca  
atagaggtgc  
1381 cccaaactcac tgacctaagc tttgttgata taaccgattc aagcatcgcc  
ctgagggtaa  
1441 ccccgctaaa ctcttccacc attattgggt accgcacac agtagttgcgc  
gcaggagaag  
1501 gtatccctat ttttgaagat tttgtggact cctcagtagg atactacaca  
30 gtcacagggc  
1561 tggagccggg cattgactat gatatcagcg ttatcactct cattaatggc  
ggcgagagtg  
1621 cccctactac actgacacaa caaacgggtg aattttgaaa acttctgcgt  
ttgagacata  
1681 gatgggttttgcatgctgcca ccagttactc cggttaaata tggatgtttc  
atgggggaag  
1741 tcagcaatttgc gccaaggatt cagataggtg gaattggggg gataaggaat  
caaatgcatc

1801 tgctaaactg attggagaaa aacacatgca atatcttcag tacactctca  
tttaaaccac  
1861 aagtagatataaaggctaga gaaatacaga tgtctgctct gttaaatata  
aaatagcaaa  
5 1921 tgttcattca atttgaagac ctagaatttt tcttcttaaa taccaaacac  
gaataccaaa  
1981 ttgcgttaagt accaattgat aagaatataat caccaaaatg taccatcatg  
ctcttccttc  
2041 taccctttga taaaactctac catgtcctt cttttagtctt aaaaacccat  
10 caaaaatttag  
2101 ggttagagtgg atgggcattt gttttaggtt ggagaaaaagt aaacttggga  
ccattctagg  
2161 ttttggct gtcacttaggt aaagaaaacac ctcttaacc acagtctggg  
gacaagcatg  
15 2221 caacattta aaggttctct gctgtgcattt ggaaaagaaa catgtgaga  
accaatttgc.  
2281 atgaacatgt tcacttgtaa gttagattca ctgaatggaa ctgttagctct  
agatatctca  
2341 catggggggaa agtttaggac cctcttgc ttttgcgtgt gtgcattgtat  
20 ttctttgtaa  
2401 agtactgcta tgtttcttt tgctgtgtgg caacttaaggc ctcttcggcc  
tgggataaaa  
2461 taatctgcag tggattttat aatgtacata aagtcaacat atttggaaagt  
agattaaaat  
25 2521 ctttttaaa tatataatg atggaaaaaa gttttttttt ggcctaaacag  
tactgtgtgt  
2581 agtgttttat tttaacagt agtacactat aactttttttt agacttagat  
tagactgttt  
2641 gcatgattat gattctgttt cctttatgca tggaaatattt attttacctt  
30 tccagctact  
2701 tcgttagttt taattttaaa atacattaaatg tgagtcttcc ttcttgc  
aaaccagctg  
2761 ttccctccctcc cactgacccgtt cgattcacca acattggtcc agacaccatg  
cgtgtcacct  
35 2821 ggg  
//

CEA  
LOCUS NM\_004363 2974 bp mRNA linear  
PRI 28-NOV-2000  
DEFINITION Homo sapiens carcinoembryonic antigen-related cell  
5 adhesion  
molecule 5 (CEACAM5), mRNA.  
ACCESSION NM\_004363  
VERSION NM\_004363.1 GI:11386170  
KEYWORDS .  
10 SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
15 Homo.  
REFERENCE 1 (bases 1 to 2974)  
AUTHORS Oikawa S, Nakazato H and Kosaki G.  
TITLE Primary structure of human carcinoembryonic antigen  
(CEA) deduced  
20 from cDNA sequence  
JOURNAL Biochem. Biophys. Res. Commun. 142 (2), 511-518 (1987)  
MEDLINE 87128144  
PUBMED 3814146  
REFERENCE 2 (bases 1 to 2974)  
25 AUTHORS Zimmermann W, Weber B, Ortlieb B, Rudert F, Schempp W,  
Fiebig HH,  
Shively JE, von Kleist S and Thompson JA.  
TITLE Chromosomal localization of the carcinoembryonic  
antigen gene  
30 family and differential expression in various tumors  
JOURNAL Cancer Res. 48 (9), 2550-2554 (1988)  
MEDLINE 88184584  
PUBMED 3356015  
REFERENCE 3 (bases 1 to 2974)  
35 AUTHORS Barnett, T., Goebel, S.J., Nothdurft, M.A. and  
Elting, J.J.  
TITLE Carcinoembryonic antigen family: characterization of  
cDNAs coding

for NCA and CEA and suggestion of nonrandom sequence variation in  
their conserved loop-domains

5 JOURNAL Genomics 3 (1), 59-66 (1988)

MEDLINE 89122014

REFERENCE 4 (bases 1 to 2974)

AUTHORS Barnett T and Zimmermann W.

TITLE Workshop report: proposed nomenclature for the  
carcinoembryonic

10 antigen (CEA) gene family

JOURNAL Tumour Biol. 11 (1-2), 59-63 (1990)

MEDLINE 90176333

PUBMED 2309067

REFERENCE 5 (bases 1 to 2974)

15 AUTHORS Schrewe H, Thompson J, Bona M, Hefta LJ, Maruya A,  
Hassauer M,  
Shively JE, von Kleist S and Zimmermann W.

TITLE Cloning of the complete gene for carcinoembryonic  
antigen: analysis

20 of its promoter indicates a region conveying cell  
type-specific

expression

JOURNAL Mol. Cell. Biol. 10 (6), 2738-2748 (1990)

MEDLINE 90258861

25 PUBMED 2342461

COMMENT PROVISIONAL REFSEQ: This record has not yet been  
subject to final  
NCBI review. The reference sequence was derived from  
M29540.1.

30 FEATURES Location/Qualifiers

source 1..2974

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="19"

/map="19q13.1-q13.2"

35 gene 1..2974

/gene="CEACAM5"

/note="CEA; CD66e"

/db\_xref="LocusID:1048"

/db\_xref="MIM:114890"  
 CDS 115..2223  
 /gene="CEACAM5"  
 /note="carcinoembryonic antigen"  
 5 /codon\_start=1  
 /db\_xref="LocusID:1048"  
 /db\_xref="MIM:114890"  
 /product="carcinoembryonic antigen-related  
 cell adhesion molecule 5"  
 10 /protein\_id="NP\_004354.1"  
 /db\_xref="GI:11386171"  
  
 /translation="MESPSAPPHRWCIPWQRLLLNASLLTFWNPPTAKLTIESTPFN  
 15 VAEGKEVLLLVHNLQPHLFGYSWYKGERVDGQRQIIGYVIGTQQATPGPAYSGREIIY  
 PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNNSKPVEDK  
 DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTFNVTRNDTASYKCETQ  
 NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGT  
 20 FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRRTVTTITVYAEPPKPFITSNNSNP  
 VEDEDALVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLGVTRNDVGPYE  
 CGIQNELSVHDSPVILNVLYGPDDPTISPSTYYPGPNLQLSCHAASNPPAQYSWL  
 25 IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPKPSISSN  
 NSKPVEDKDAVAFTEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTFNVTRND  
 RAYVCGIQNSVSANRSDPVTLVLYGPDPTIISPPDSSYLSGANLNLSCHSASNPSQ  
 YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNNSIVKSITVSASGTSPG  
 30 LSGATVGIMIGVLVGVALL"  
 misc\_feature 595..780  
 /note="IG; Region: Immunoglobulin"  
 misc\_feature 868..1026  
 35 /note="IG; Region: Immunoglobulin"  
 misc\_feature 1129..1314  
 /note="IG; Region: Immunoglobulin"  
 misc\_feature 1399..1566  
 /note="IG; Region: Immunoglobulin"  
 misc\_feature 1930..2091  
 35 /note="IG; Region: Immunoglobulin"  
  
 BASE COUNT 840 a 847 c 613 g 674 t  
 ORIGIN

1 ctcagggcag agggaggaag gacagcagac cagacagtca cagcagcctt  
gacaaaacgt  
61 tcctggaact caagctttc tccacagagg aggacagagc agacagcaga  
gaccatggag  
5 121 tctccctcgg cccctccccca cagatggtgc atcccctggc agaggctcct  
gctcacagcc  
181 tcacttctaa ccttctggaa cccgcccacc actgccaagc tcactattga  
atccacgcgg  
241 ttcaatgtcg cagagggaa ggaggtgctt ctacttgtcc acaatctgcc  
10 ccagcatctt  
301 ttggctaca gctggtacaa aggtgaaaga gtggatggca accgtcaaatt  
tataggatat  
361 gtaataggaa ctcaacaagc taccggcaggg cccgcataca gtggtcgaga  
gataatatac  
15 421 cccaatgcat ccctgctgat ccagaacatc atccagaatg acacaggatt  
ctacaccctt  
481 cacgtcataa agtcagatct tgtgaatgaa gaagcaactg gccagttccg  
ggtatacccg  
541 gagctgccccca agccctccat ctccagcaac aactccaaac ccgtggagga  
20 20  
caaggatgct  
601 gtggccttca cctgtgaacc tgagactcag gacgcaacct acctgtggtg  
ggtaaacaat  
661 cagagcctcc cggtcagtcc caggctgcag ctgtccaaatg gcaacaggac  
cctcactcta  
25 721 ttcaatgtca caagaaatga cacagcaagc tacaaatgtg aaacccagaa  
cccagtgagt  
781 gccaggcgca gtgattcagt catcctgaat gtcctctatg gcccggatgc  
ccccaccatt  
841 tccctctaa acacatctt cagatcaggg gaaaatctga acctctcctg  
30 ccacgcagcc  
901 tctaaccac ctgcacagta ctcttggttt gtcaatggga ctttcagca  
atccacccaa  
961 gagctttta tcccaacat cactgtgaat aatagtggat cctatacgtg  
ccaagcccat  
35 1021 aactcagaca ctggcctcaa taggaccaca gtcacgcacgca tcacagtcta  
tgctgagcc  
1081 cccaaaccct tcattcaccag caacaactcc aaccccggtgg aggtgagga  
tgctgttagcc

1141 ttaacctgtg aacctgagat tcagaacaca acctacctgt ggtgggtaaa  
taatcagac  
1201 ctcccggtca gtcccaggct gcagctgtcc aatgacaaca ggaccctcac  
tctactcagt  
5 1261 gtcacaagga atgatgttagg accctatgag tgtgaaatcc agaacgaatt  
aagtgttgac  
1321 cacagcgacc cagtcatcct gaatgtcctc tatggccag acgaccccac  
catttcccc  
1381 tcatacacctt attaccgtcc aggggtgaac ctcagccctc cctgccatgc  
10 agcctctaac  
1441 ccacctgcac agtattcttg gctgattgat gggAACatcc agcaacacac  
acaagagctc  
1501 tttatctcca acatcactga gaagaacagc ggactctata cctgccaggc  
caataactca  
15 1561 gccagtggcc acagcaggac tacagtcaag acaatcacag tctctgcgg  
gctgccc  
1621 ccctccatct ccagcaacaa ctccaaaccc gtggaggaca agatgctgt  
ggccttcacc  
1681 tgtgaacctg aggctcagaa cacaacctac ctgtggtggg taaatggtca  
20 gaggcctccca  
1741 gtcagtccca ggctgcagct gtccaatggc aacaggaccc tcactctatt  
caatgtcaca  
1801 agaaatgacg caagagccta tgtatgtgga atccagaact cagtgagtgc  
aaaccgcagt  
25 1861 gaccaggatca ccctggatgt cctctatggg ccggacaccc ccatcatttc  
ccccccagac  
1921 tcgtcttacc ttccggagc gaacctcaac ctctcctgcc actccggcctc  
taaccatcc  
1981 ccgcagtatt cttggcgtat caatggata ccgcagcaac acacacaagt  
30 tctcttatac  
2041 gccaaaatca cgccaaataa taacgggacc tatgcctgtt ttgtctctaa  
cttggctact  
2101 gcccgcata attccatagt caagagcatc acagtctctg catctggAAC  
ttctcctgg  
35 2161 ctctcagctg gggccactgt cggcatcatg attggagtgc tggttgggg  
tgctctgata  
2221 tagcagccct ggtgttagttt cttcatttca ggaagactga cagttgtttt  
gcttcttcct

```

2281 taaaggcattt gcaacagcta cagtctaaaa ttgcttcttt accaaggata
tttacagaaa
2341 agactctgac cagagatcga gaccatccta gccaacatcg taaaacccca
tctctactaa
5 2401 aaatacataaa atgagctggg cttgggtggcg cgcacctgta gtcccagttt
ctcgaggc
2461 tgaggcagga gaatcgctt aacccgggag gtggagattg cagttagcccc
agatcgacc
2521 actgcactcc agtctggcaa cagagcaaga ctccatctca aaaagaaaaag
10 aaaagaagac
2581 tctgacctgt actcttgaat acaagtttct gataccactg cactgtctga
gaatttccaa
2641 aacttaatg aactaactga cagcttcatg aaactgtcca ccaagatcaa
gcagagaaaa
15 2701 taattaattt catggacta aatgaactaa tgaggattgc tgattctta
aatgtcttgt
2761 ttccagatt tcagggaaact tttttcttt taagctatcc actcttacag
caatttgata
2821 aaatataactt ttgtgaacaa aaattgagac atttacattt tctccatag
20 tggtcgtcc
2881 agacttggga aactattcat gaatatttattt attgtatggt aatatagtta
ttgcacaagt
2941 tcaataaaaaa tctgctcttt gtataacaga aaaa
//
```

25

**Her2/Neu**

```

LOCUS HUMHER2A 4530 bp mRNA linear
PRI 18-SEP-1995
30 DEFINITION Human tyrosine kinase-type receptor (HER2) mRNA,
complete cds.
ACCESSION M11730
VERSION M11730.1 GI:183986
KEYWORDS tyrosine kinase.
35 SOURCE Homo sapiens (clone: lambda-HER2-436) fetal cDNA to
mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
 Homo.  
 REFERENCE 1 (bases 1 to 4530)  
 AUTHORS Coussens, L., Yang-Feng, T.L., Liao, Y.-C., Chen, E.,  
 5 Gray, A., McGrath, J., Seeburg, P.H., Libermann, T.A.,  
 Schlessinger, J., Francke, U., Levinson, A. and Ullrich, A.  
 TITLE Tyrosine kinase receptor with extensive homology to  
 10 EGF receptor shares chromosomal location with neu oncogene  
 JOURNAL Science 230 (4730), 1132-1139 (1985)  
 MEDLINE 86070181  
 REFERENCE 2 (bases 1701 to 1719)  
 15 AUTHORS Ullrich, A.  
 JOURNAL Unpublished (1988)  
 FEATURES Location/Qualifiers  
 source 1..4530  
 /organism="Homo sapiens"  
 20 /db\_xref="taxon:9606"  
 /clone="lambda-HER2-436"  
 /dev\_stage="fetal"  
 mRNA <1..4530  
 /product="HER2 mRNA"  
 25 CDS 151..3918  
 /note="HER2 receptor"  
 /codon\_start=1  
 /protein\_id="AAA75493.1"  
 /db\_xref="GI:306840"  
 30 /translation="MELAALCRWGLLALLPPGAASTQVCTGTDMLRLPASPETHLD  
 MLRHLYQGCQVVQGNLELTLYLPTNASLSQLQDIQEVGQGYVLIAHNQVRQVPLQRLRIV  
 RGTQLFEDNYALAVLDNGDPLNNTPVTGASPGLRELQLRSLTEILKGGVLIQRNPQ  
 LCYQDTILWKDIFHKNNQLALTLIDTNRSRACHPCSPMCKGSRCWGESSEDCQSLTRT  
 35 VCAGGCCARCKGPLTDCCHCQECAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNT  
 DTFESMPNPEGRYTFGASCVTACPYNLSTDVGSCTLVCPLHNQEVTAEDEGTQRCEKC  
 SKPCARVCYGLGMEHLREVRAVTSANIQEAGCKKIFGSLAFLPESFDGDPASNTAPL  
 QPEQLQVFETLEEITGYLYISAWPDSLPDLSVFQNLQVIRGRILHNGAYSILTLQGLGI  
 SWLGLRSLRELGSGLALIHNNTHLCFVHTVPWDQLFRNPHQALLHTANRPEDECVGEG

LACHQLCARGHCWGPPTQCVNCSQFLRGQECVEECRVLQGLPREYVNARHCLPCHPE  
 CQPQNGSVTCFGPEADQCVACAHYKDPPFCVARCPSGVKPDLSSYMPIWKFPDEEGACQ  
 PCPINCTHSCVDLDDKGCPAEQRASPLTSIVSAVVGILLVVVLGVVFGILIKRRQQKI  
 RKYTMRLLQETELVEPLTPSGAMPNQAQMRLKETELRKVKVLGSGAFGTYYKGIWI  
 5 PDGENVKIPVAKVLRENTSPKANKEILDEAYVMAGVGSPYVSRLLGICLTSTVQLVT  
 QLMPYGCCLLDHVRENRRGRLGSQDLLNWCMQIAKGMSTYLEDVRLVHDLAARNVLVKSP  
 NHVKITDFGLARLLDIDETEYHADGGKVIKWMALESILRRRFTHQSDVWSYGVTVWE  
 LMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYMIMVKCWMIDSECRPRFREL  
 VSEFSRMARDPQRFVVIQNEDELDGPASPLDSTFYRSLLLEDDDMGDLVDAEYLVPPQQGF  
 10 FCPDPAPGAGGMVHHRHRSSTRSGGGDLTLGLEPSEEAPRSP LAPSEGAGSDVFDG  
 DLGMGAAKGLQSLPLTHDPSPQLQRYSEDPTVPLPSETDGYVAPLTCS PQPEYVNQPDVR  
 PQPPSPREGPLPAARPAGATLERAKTLSPGKNGVVKDVFAGGAAVENPEYLTPQGGAA  
 PQPHPPPAAFSPAFDNLYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV"  
 old\_sequence 1701..1719  
 15 /citation=[1]  
 BASE COUNT 922 a 1382 c 1346 g 880 t  
 ORIGIN Chromosome 17q21-q22.  
 1 aattctcgag ctcgtcgacc ggtcgacgag ctcgagggtc gacgagctcg  
 agggcgcgcg  
 20 61 cccggccccc acccctcgca gcaccccgcg ccccgccccc tcccagccgg  
 gtccagccgg  
 121 agccatgggg ccggagccgc agtgagcacc atggagctgg cggccttgtg  
 ccgctgggggg  
 181 ctcccttcg ccctttgcc ccccgagcc gcgagcaccc aagtgtgcac  
 25 cggcacagac  
 241 atgaagctgc ggctccctgc cagtcccgag acccacctgg acatgctccg  
 ccacctctac  
 301 cagggctgcc aggtggtgca gggaaacctg gaactcacct acctgcccac  
 caatgccagc  
 361 ctgtccttcc tgcaggatat ccaggagggtg cagggctacg tgctcatcgc  
 tcacaaccaa  
 421 gtgaggcagg tcccactgca gaggctgcgg attgtgcgag gcacccagct  
 ctttggggac  
 481 aactatgccc tggccgtgct agacaatgga gaccgcgtga acaataaccac  
 35 ccctgtcaca  
 541 ggggcctccc caggaggct gcgggagctg cagttcgaa gcctcacaga  
 gatcttgaaa  
 601 ggaggggtct tgcgtcccg gaaccccccag ctctgttacc aggacacgat  
 tttgtggaaag

661 gacatcttcc acaagaacaa ccagctggct ctcacactga tagacaccaa  
ccgctctcg  
721 gcctgccacc cctgttctcc gatgtgtaaag ggctcccgct gctggggaga  
gagttctgag  
5 781 gattgtcaga gcctgacgctg cactgtctgt gccggtggct gtgcccgtg  
caaggggcca  
841 ctgcccactg actgctgcca tgagcagtgt gctgccggct gcacgggccc  
caagcactct  
901 gactgcctgg cctgcctcca cttcaaccac agtggcatct gtgagctgca  
10 ctgcccagcc  
961 ctggtcacct acaacacaga cacgttttag tccatgccc atcccggagg  
ccggtataca  
1021 ttccggcgcca gctgtgtgac tgcctgtccc tacaactacc tttctacgga  
cgtgggatcc  
15 1081 tgcaccctcg tctgccccct gcacaaccaa gaggtgacag cagaggatgg  
aacacagccg  
1141 tgtgagaagt gcagcaagcc ctgtgcccga gtgtgctatg gtctggcat  
ggagcacttg  
1201 cgagaggtga gggcagttac cagtgccaat atccaggagt ttgctggctg  
20 caagaagatc  
1261 tttgggagcc tggcattct gccggagagc tttgatgggg acccagcctc  
caacactgcc  
1321 ccgctccagc cagagcagct ccaagtgttt gagactctgg aagagatcac  
aggtaaccta  
25 1381 tacatctcag catggccgga cagcctgcct gacccatcg tttccagaa  
cctgcaagta  
1441 atccggggac gaattctgca caatggcgcc tactcgctga ccctgcaagg  
gctgggcatc  
1501 agctggctgg ggctgcgcctc actgagggaa ctgggcagtg gactggccct  
30 catccaccat  
1561 aacacccacc tctgcttcgt gcacacggtg ccctggacc agctcttcg  
gaacccgcac  
1621 caagctctgc tccacactgc caaccggcca gaggacgagt gtgtggcgaa  
gggcctggcc  
35 1681 tgccaccagc tgtgcgcctg agggcactgc tgggtccag ggcccacccaa  
gtgtgtcaac  
1741 tgcagccagt tccttcgggg ccaggagtgc gtggaggaat gccgagtaact  
gcaggggctc

1801 cccaggaggt atgtgaatgc caggcactgt ttgccgtgcc accctgagtg  
tcagccccag  
1861 aatggctcag tgacctgttt tggaccggag gctgaccagt gtgtggcctg  
tgcacctat  
5 1921 aaggaccctc ccttcgttgt ggccccgtgc cccagcggtg taaaacctga  
ccttccttac  
1981 atgcccattct ggaagttcc agatgaggag ggccgtatgcc agccttgc  
catcaactgc  
2041 acccactctt gtgtggacct ggatgacaag ggctgccccg cccagcgag  
10 agccagccct  
2101 ctgacgttcca tcttcgttgt ggtgggtggc attctgttgtt tctgtggctt  
gggggtggtc  
2161 tttgggatcc tcatcaagcg acggcagcag aagatccgga agtacacgt  
gcggagactg  
15 2221 ctgcaggaaa cggagcttgt ggagccgtg acacctagcg gagcgatgcc  
caaccaggcg  
2281 cagatgcgga tcctgaaaga gacggagctg aggaagggtga aggtgcttgg  
atctggcgct  
2341 tttggcacag tctacaaggg catctggatc cctgtatgggg agaatgtgaa  
20 aattccagtg  
2401 gccatcaaag ttttgggaa aaacacatcc cccaaagcca acaaagaaaat  
cttagacgaa  
2461 gcatacgtga tggctggtgt gggctccca tatgtctccc gccttcgtgg  
catctgcctg  
25 2521 acatccacgg tgcagcttgt gacacagctt atgcctatg gtcgccttt  
agaccatgtc  
2581 cggaaaaacc gcgacgcct gggctccca gacctgtga actggtgtat  
gcagattgcc  
2641 aaggggatga gtcaccttgg gatgtgcgg ctcgtacaca gggacttggc  
30 cgctcgaaac  
2701 gtgctggtca agagtccaa ccatgtcaaa attacagact tcggctggc  
tcggctgtg  
2761 gacatttgcg agacagagta ccatgcagat gggggcaagg tgccatcaa  
gtggatggcg  
35 2821 ctggagtcca ttctccgcg gcggttcacc caccagagt atgtgtggag  
ttatggtgtg  
2881 actgtgtggg agctgtatgac ttttggggcc aaaccttacg atggatccc  
agcccgaggag

2941 atccctgacc tgctggaaaa gggggagcgg ctgccccagc ccccccacatctg  
caccattgtat  
3001 gtctacatga tcatggtcaa atgttggatg attgactctg aatgtcggcc  
aagattccgg  
5 3061 gagttgggtgt ctgaattctc ccgcattggcc agggacccccc agcgctttgt  
ggtcattccag  
3121 aatgaggact tggggccagc cagtccttgc gacagcacct tctaccgctc  
actgctggag  
3181 gacgatgaca tgggggacct ggtggatgct gaggagtac tggtaaaaa  
10 gcagggttc  
3241 ttctgtccag accctgcccc gggcgctggg ggcattggcc accacaggca  
ccgcagctca  
3301 tctaccagga gtggcggtgg ggacctgaca cttagggctgg agccctctga  
agaggaggcc  
15 3361 cccaggtctc cactggcacc ctccgaaggg gctggctccg atgtatttga  
tggtgacctg  
3421 ggaatggggg cagccaaagg gctgcaaagc ctccccacac atgacccca  
ccctctacag  
3481 cggtacagtg aggacccac agtacccctg ccctctgaga ctgatggcta  
20 cgttgcccc  
3541 ctgacctgca gccccagcc tgaatatgtg aaccagccag atgttcggcc  
ccagccccct  
3601 tcgccccgag agggccctct gcctgctgcc cgacctgctg gtgccactct  
ggaaagggcc  
25 3661 aagactctct ccccaggaa gaatgggtc gtcaaagacg tttttgcctt  
tgggggtgcc  
3721 gtggagaacc ccgagttactt gacacccag ggaggagctg cccctcagcc  
ccaccctctt  
3781 cctgccttca gcccagccctt cgacaacctc tattactggg accaggaccc  
30 accagagcgg  
3841 ggggctccac ccagcacctt caaaggaca cctacggcag agaacccaga  
gtacctgggt  
3901 ctggacgtgc cagtgtgaac cagaaggcca agtccgcaga agccctgtatg  
tgtcctcagg  
35 3961 gagcaggaa ggcctgactt ctgctggcat caagaggtgg gaggggccctc  
cgaccacttc  
4021 caggggaacc tgccatgcca ggaacctgtc ctaaggaacc ttcccttcctg  
cttgagttcc

```

        4081 cagatggctg gaaggggtcc agcctcgttg gaagaggaac agcactgggg
        agtctttgtg
        4141 gattctgagg ccctgccaa tgagactcta gggtccagtg gatgccacag
        cccagcttgg
5       4201 cccttcctt ccagatcctg ggtactgaaa gccttaggga agctggcctg
        agaggggaag
        4261 cggccctaag ggagtgtcta agaacaaaag cgaccattc agagactgtc
        cctgaaacct
        4321 agtactgccc cccatgagga aggaacagca atggtgtcag tatccaggct
10      ttgtacagag
        4381 tgctttctg tttagttttt acttttttt ttttgtttt ttaaagacga
        aataaaagacc
        4441 caggggagaa tgggtgttgt atggggaggc aagtgtgggg ggtccttctc
        cacacccact
15      4501 ttgtccattt gcaaataatat tttggaaaac
        //

```

```

SCP-1
20     LOCUS          HSSCP1PRT                      3393 bp      mRNA      linear
        PRI 26-FEB-1997
        DEFINITION  H.sapiens mRNA for SCP1 protein.
        ACCESSION   X95654
        VERSION     X95654.1  GI:1212982
25     KEYWORDS    SCP1 gene.
        SOURCE       human.
        ORGANISM    Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
                    Euteleostomi;
        30       Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
                    Homo.
        REFERENCE   1 (bases 1 to 3393)
        AUTHORS     Meuwissen,R.L., Meerts,I., Hoovers,J.M., Leschot,N.J.
                    and
        35       Heyting,C.
        TITLE       Human synaptonemal complex protein 1 (SCP1): isolation
                    and
                    characterization of the cDNA and chromosomal
                    localization of the

```

gene  
 JOURNAL Genomics 39 (3), 377-384 (1997)  
 MEDLINE 97224467  
 REFERENCE 2 (bases 1 to 3393)  
 5 AUTHORS Meuwissen, R.J.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-FEB-1996) Dr. R.L.J. Meuwissen,  
 Agricultural  
 University, Genetics, Dreijenlaan 2, 6703 HA  
 10 WAGENINGEN,  
 NETHERLANDS  
 FEATURES Location/Qualifiers  
 source 1..3393  
 /organism="Homo sapiens"  
 15 /db\_xref="taxon:9606"  
 /map="1p12-p13"  
 /tissue\_type="testis"  
 gene 95..3025  
 /gene="SCP1"  
 20 CDS 95..3025  
 /gene="SCP1"  
 /codon\_start=1  
 /evidence=experimental  
 /product="polypeptide of 976 aa"  
 25 /protein\_id="CAA64956.1"  
 /db\_xref="GI:1212983"  
 /db\_xref="SWISS-PROT:Q15431"  
  
 /translation="MEKQKPFALFVPPRSSSSQVS AVKPQTLGGDSTFFKSFNKCTED  
 30 DLEFPFAKTNLSKNGENIDSDPALQKVNFPLVLEQVGNSDCHYQEGLKDSLENSEGL  
 SRVFSKLYKEAEKIKKWKVSTEAE LRQKESKLQENRKIIEAQRKAIQELQFGNEKVSL  
 KLEEGIQENKDLIKENNATRHLCNLLKETCARS A EKTKKYEYEREETRQVYMDLNNNI  
 EKMITAHGELRVQAENSRL EMHFKLKEDYEKIQHLEQ EYKKEINDKEKVSL LLIQIT  
 EKENKMKDLTFLLEESRDKVNLQLEEKTKLQSENLKQSIEKQHHLTKELEDIKVSLQRS  
 35 VSTQKALEEDLQIATKTICQLTEEKETQMEESNKARA AHSFVVTEFETTVCSLEELLR  
 TEQQRLEKNEDQLKILTMELQKKSSELEEMTKLTNNKEVELEELKKVLGEKETLLYEN  
 KQFEKIAEELKGTEQELIGLLQAREKEVHDLEIQLTAITTSEQYSKEVKDLKTELEN  
 EKLKNTELTSHCNKLSLENKELTQETSDMTLELKNOQEDINNNKKQEEERMLKQIENLQ  
 ETETQLRNELEYVREELKQRDEVKCKLDKSEENCNLRKQVENKNKYIEELQQENKA"

LKKKGTAESKQLNVYEIKVNKLELELESQKQKFGEITDTYQKEIEDKKISEENLLEEV  
 EKAKVIADEAVKLQKEIDKRCQHKIAEMVALMEKKHQYDKIIERDSELGLYKSKEQ  
 EQSSLRASLEIELSNLKAELLSVKKQLEIEREKEKLREAKENTATLKEKKDKKTQ  
 5 FLLETPEIYWKLDASKAVPSQTVSRNFTSDHGISKDKRDYLWTSAKNTLSTPLPKAYT  
 VKTPTKPKLQQRENLNIPIEESKKRKMAFEFDINSDSSETDLSMVSEETLKTLY  
 RNNNPPASHLCVTPKKAPSSLTPGPTLKFGAIRKMRDRWAVIAKMDRKKKLKEAE  
 KLFV"  
 BASE COUNT 1440 a 483 c 648 g 822 t  
 ORIGIN  
 10 1 gcccctcatag accgtttgtt gtagttcgcg tgggaacacgc aacccacggc  
 ttcccgatag  
 61 ttcttcaaag atatttacaa ccgtaacaga gaaaatggaa aagcaaaagc  
 cctttgcatt  
 121 gttcgtacca ccgagatcaa gcagcagtca ggtgtctcg gtgaaacctc  
 15 agaccctggg  
 181 aggcgattcc actttttca agagtttcaa caaatgtact gaagatgatt  
 tggagtttcc  
 241 atttgcaaag actaatctct ccaaaaatgg ggaaaacatt gattcagatc  
 ctgtcttaca  
 20 301 aaaagttaat ttcttgcgc tgcttgagca ggttggtaat tctgactgtc  
 actatcagga  
 361 aggactaaaa gactctgatt tggagaattc agagggattt agcagagtgt  
 tttcaaaact  
 421 gtataaggag gctgaaaaga taaaaaaatg gaaagtaagt acagaagctg  
 25 aactgagaca  
 481 gaaagaaaatg aagttgcaag aaaacagaaaa gataattgaa gcacagcgaa  
 aagccattca  
 541 ggaactgcaa tttggaaatg aaaaagtaag tttgaaatta gaagaaggaa  
 tacaagaaaa  
 30 601 taaagattta ataaaagaga ataatgccac aaggcattta tgtaatctac  
 tcaaagaaaa  
 661 ctgtgctaga tctgcagaaaa agacaaaagaa atatgaatat gaacggaaag  
 aaaccaggca  
 721 agtttatatg gatctaaata ataacattga gaaaatgata acagctcatg  
 35 gggaaacttcg  
 781 tgtgcaagct gagaattcca gactggaaat gcatttaag ttaaaggaag  
 attatgaaaa  
 841 aatccaacac cttgaacaag aatacaagaa ggaaataaat gacaaggaaa  
 agcaggtatc

901 actactattg atccaaatca ctgagaaaaga aaataaaatg aaagatttaa  
catttctgct  
961 agaggaatcc agagataaag ttaatcaatt agagggaaaag acaaaaattac  
agagtggaaaa  
5 1021 cttaaaaacaa tcaattgaga aacagcatca tttgactaaa gaactagaag  
atattaaagt  
1081 gtcattacaa agaagtgtga gtactcaaaa ggcttttagag gaagatttac  
agatagcaac  
1141 aaaaacaatt tgcagctaa ctgaagaaaa agaaaactcaa atggaagaat  
10 ctaataaagc  
1201 tagagctgct cattcgtttgc tggttactga atttggaaact actgtctgca  
gcttggaaaga  
1261 attattgaga acagaacagc aaagatttggaa aaaaaatgaa gatcaattgaa  
aaataacttac  
15 1321 catggagctt caaaagaaat caagtgagct ggaagagatg actaagctta  
caaataacaa  
1381 agaagtagaa cttgaagaat tgaaaaaaagt cttgggagaa aaggaaacac  
ttttatatga  
1441 aaataaaacaa tttgagaaga ttgctgaaga attaaaagga acagaacaag  
20 aactaattgg  
1501 tcttctccaa gccagagaga aagaagtaca tgatttggaa atacagttaa  
ctgccattac  
1561 cacaagtgaa cagtattatt caaaagaggt taaagatcta aaaactgagc  
ttgaaaacgaa  
1621 gaagcttaag aatactgaat taacttcaca ctgcaacaag ctttcactag  
aaaacaaaga  
1681 gctcacacag gaaacaagtg atatgaccct agaactcaag aatcagcaag  
aagatattaa  
1741 taataacaaa aagcaagaag aaaggatgtt gaaacaaata gaaaatcttc  
30 aagaaacaga  
1801 aacccaaatta agaaatgaac tagaatatgt gagagaagag ctaaaacaga  
aaagagatga  
1861 agttaaatgt aaattggaca agagtgaaga aaattgtac aatttaagga  
aacaagttga  
35 1921 aaataaaaac aagtatattg aagaacttca gcaggagaat aaggcccttga  
aaaaaaaagg  
1981 tacagcagaa agcaagcaac tgaatgttta tgagataaag gtcaataaat  
tagagttga

2041 actagaaaagt gccaaacaga aatttggaga aatcacagac acctatcaga  
aagaaaattga  
2101 ggacaaaaag atatcagaag aaaatctttt ggaagaggtt gagaaagcaa  
aagtaatagc  
5 2161 tcatgaagca gtaaaattac agaaagaaat tgataagcga tgtcaacata  
aaatagctga  
2221 aatggtagca cttatggaaa aacataagca ccaatatgat aagatcattg  
aagaaagaga  
10 2281 ctcagaattha ggactttata agagcaaaga acaagaacag tcatactga  
gagcatctt  
2341 ggagattgaa ctatccaatc tcaaagctga acttttgcgt gtttggaaagc  
aacttggaaat  
2401 agaaagagaa gagaaggaaa aactcaaaag agaggcaaaa gaaaacacag  
ctactcttaa  
15 2461 agaaaaaaaaa gacaagaaaa cacaacatt tttattggaa acacctgaaa  
tttattggaa  
2521 attggattct aaagcagttc ctccacaaac tgtatctga aatttcacat  
cagttgatca  
2581 tggcatatcc aaagataaaa gagactatct gtggacatct gccaaaata  
20 ctttatctac  
2641 accattgcca aaggcatata cagtgaagac accaacaaaaa ccaaaactac  
agcaagaga  
2701 aaacttgaat ataccattt aagaaagtaa aaaaaagaga aaaatggcct  
ttgaatttga  
2761 tattaattca gatagttcag aaactactga tctttgagc atggtttcag  
aagaagagac  
2821 attgaaaaca ctgtatagga acaataatcc accagcttc catcttgc  
tcaaaacacc  
2881 aaaaaaggcc cttcatctc taacaacccc tggacctaca ctgaagtttgc  
30 gagctataag  
2941 aaaaatgcgg gaggaccgtt gggctgtaat tgctaaaatg gatagaaaaa  
aaaaactaaa  
3001 agaagctgaa aagttatttgc ttaatttca gagaatcagt gtagttaagg  
agcctaataa  
35 3061 cgtgaaactt atagttata ttttgcgtt atttgccaga gccacattt  
atctggaagt  
3121 tgagacttaa aaaatacttg catgaatgat ttgtgtttt tttatatttt  
agcctaataatg

3181 ttaactacat attgtctgga aacctgtcat tgtattcaga taatttagatg  
attatatatt  
3241 gttgttactt tttcttgtat tcatgaaaac tgttttact aagtttcaa  
atttgtaaag  
5 3301 ttagccttg aatgcttagga atgcattatt gagggcatt ctttattctt  
tactattaaa  
3361 atatttgga tgcaaaaaaaaaaaaaaaa aaa  
//

10

**SSX-4**

LOCUS NM\_005636 576 bp mRNA linear  
PRI 10-DEC-2001  
DEFINITION Homo sapiens synovial sarcoma, X breakpoint 4 (SSX4),  
15 mRNA.  
ACCESSION NM\_005636  
VERSION NM\_005636.1 GI:5032122  
KEYWORDS .  
SOURCE human.  
20 ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
Homo.  
25 REFERENCE 1 (bases 1 to 576)  
AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S.,  
Scanlan,M.J., Knuth,A.,  
Pfreundschuh,M., Old,L.J. and Chen,Y.T.  
TITLE SSX: a multigene family with several members  
30 transcribed in normal  
testis and human cancer  
JOURNAL Int. J. Cancer 72 (6), 965-971 (1997)  
MEDLINE 98021352  
COMMENT PROVISIONAL REFSEQ: This record has not yet been  
35 subject to final  
NCBI review. The reference sequence was derived from  
U90841.1.  
FEATURES Location/Qualifiers  
source 1..576

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.3"
5      gene      1..576
/gene="SSX4"
/db_xref="LocusID:6759"
/db_xref="MIM:300326"
CDS      1..567
10     gene      /gene="SSX4"
/note="Kruppel-associated box containing SSX
gene"
/codon_start=1
/db_xref="LocusID:6759"
15     /db_xref="MIM:300326"
/product="synovial sarcoma, X breakpoint 4"
/protein_id="NP_005627.1"
/db_xref="GI:5032123"

20     /translation="MNGDDAFARRPRDDAQISEKLRKAFDDIAKYFSKKEWEEKMSSEKIVY
VYMKLNYEVMTKLGFKVLPPFMRSKRAADFHGNDFGNDRNHRNQVERPQMTFG
SLQRIFPKIMPKKPAEEENGLKEVPEASGPQNDGKQLCPPGNPSTLEKINKTSGPKRG
KHAWTHRLRERKQLVYEEISDPBEDDE"
misc_feature 70..246
25     /note="KRAB; Region: krueppel associated box"
BASE COUNT      187 a      127 c      150 g      112 t
ORIGIN
1      atgaacggag acgacgcctt tgcaaggaga cccagggatg atgctcaaat
atcagagaag
30     61 ttacgaaagg cttcgatga tattgccaaa tacttctcta agaaagagtg
ggaaaaagatg
121    aaatcctcg agaaaaatgt ctatgtgtat atgaagctaa actatgaggat
catgactaaa
181    ctagtttca aggtcacccct cccaccttcc atgcgttagta aacgggctgc
35     agacttccac
241    gggaatgatt ttggtaacga tcgaaaccac aggaatcagg ttgaacgtcc
tcagatgact
301    ttcggcagcc tccagagaat cttcccgaaag atcatgcccc agaagccagc
agaggaagaa

```

361 aatggtttga aggaagtgcc agaggcatct ggcccacaaa atgatggaa  
acagctgtgc  
421 cccccggaa atccaagtac cttggagaag attaacaaga catctggacc  
caaaaggggg  
5 481 aaacatgcct ggaccacag actgcgtgag agaaagcagc tggtggttta  
tgaagagatc  
541 agcgaccctg aggaagatga cgagtaactc ccctcg

10 All patents and publications mentioned in the specification  
are indicative of the levels of those skilled in the art to which  
the invention pertains. All patents and publications are herein  
incorporated by reference to the same extent as if each individual  
publication was specifically and individually indicated to be  
incorporated by reference.

15 The invention illustratively described herein suitably may be  
practiced in the absence of any element or elements, limitation or  
limitations which is not specifically disclosed herein. The terms  
and expressions which have been employed are used as terms of  
description and not of limitation, and there is no intention that  
20 in the use of such terms and expressions indicates the exclusion  
of equivalents of the features shown and described or portions  
thereof. It is recognized that various modifications are possible  
within the scope of the invention claimed. Thus, it should be  
understood that although the present invention has been  
25 specifically disclosed by preferred embodiments and optional  
features, modification and variation of the concepts herein  
disclosed may be resorted to by those skilled in the art, and that  
such modifications and variations are considered to be within the  
scope of this invention as defined by the appended claims.

30

WHAT IS CLAIMED IS:

1. An isolated epitope, comprising a component selected from the group consisting of:
  - (i) a polypeptide having the sequence as disclosed in TABLE 1;
  - (ii) an epitope cluster comprising the polypeptide of (i);
  - (iii) a polypeptide having substantial similarity to (i) or (ii);
  - (iv) a polypeptide having functional similarity to any of (i) through (iii); and
  - (v) a nucleic acid encoding the polypeptide of any of (i) through (iv).
- 5 2. The epitope of claim 1, wherein the epitope is immunologically active.
- 10 3. The epitope of claim 1, wherein the polypeptide is less than about 30 amino acids in length.
4. The epitope of claim 1, wherein the polypeptide is 8 to 10 amino acids in length.
- 15 5. The epitope of claim 1, wherein the substantial or functional similarity comprises addition of at least one amino acid.
6. The epitope of claim 5, wherein the at least one additional amino acid is at an N-terminus of the polypeptide.
- 20 7. The epitope of claim 1, wherein the substantial or functional similarity comprises a substitution of at least one amino acid.
8. The epitope of claim 1, the polypeptide having affinity to an HLA-A2 molecule.
- 25 9. The epitope of claim 8, wherein the affinity is determined by an assay of binding.
10. The epitope of claim 8, wherein the affinity is determined by an assay of restriction of epitope recognition.
11. The epitope of claim 8, wherein the affinity is determined by a prediction algorithm.
- 30 12. The epitope of claim 1, the polypeptide having affinity to an HLA-B7 or HLA-B51 molecule.
13. The epitope of claim 1, wherein the polypeptide is a housekeeping epitope.
14. The epitope of claim 1, wherein the polypeptide corresponds to an epitope displayed on a tumor cell.
15. The epitope of claim 1, wherein the polypeptide corresponds to an epitope displayed on a neovasculature cell.
- 35 16. The epitope of claim 1, wherein the peptide is an immune epitope.
17. The epitope of claim 1 wherein the epitope is a nucleic acid.
18. A pharmaceutical composition comprising the peptide of claim 1 and a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.
19. The composition of claim 18, where the adjuvant is a polynucleotide.

20. The composition of claim 19 wherein the polynucleotide comprises a dinucleotide.
21. The composition of claim 20 wherein the dinucleotide is CpG.
22. The composition of claim 18, wherein the adjuvant is encoded by a polynucleotide.
23. The composition of claim 18 wherein the adjuvant is a cytokine.
- 5 24. The composition of claim 23 wherein the cytokine is GM-CSF.
25. The composition of claim 18 further comprising a professional antigen-presenting cell (pAPC).
  26. The composition of claim 25, wherein the pAPC is a dendritic cell.
  27. The composition of claim 18, further comprising a second epitope.
- 10 28. The composition of claim 27, wherein the second epitope is a polypeptide.
29. The composition of claim 27, wherein the second epitope is a nucleic acid.
30. The composition of claim 27, wherein the second epitope is a housekeeping epitope.
31. The composition of claim 27, wherein the second epitope is an immune epitope.
- 15 32. A pharmaceutical composition comprising the nucleic acid of claim 1 and a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.
33. A recombinant construct comprising the nucleic acid of Claim 1.
34. The construct of claim 33, further comprising a plasmid, a viral vector, or an artificial chromosome.
- 20 35. The construct of claim 33, further comprising a sequence encoding at least one feature selected from the group consisting of a second epitope, an IRES, an ISS, an NIS, and ubiquitin.
36. A purified antibody that specifically binds to the epitope of claim 1.
37. A purified antibody that specifically binds to a peptide-MHC protein complex comprising the epitope of claim 1.
- 25 38. The antibody of claim 36 or claim 37, wherein the antibody is a monoclonal antibody.
39. A multimeric MHC-peptide complex comprising the epitope of claim 1.
40. An isolated T cell expressing a T cell receptor specific for an MHC-peptide complex, the complex comprising the epitope of claim 1.
- 30 41. The T cell of claim 40, produced by an *in vitro* immunization.
42. The T cell of claim 40, isolated from an immunized animal.
43. A T cell clone comprising the T cell of claim 40.
44. A polyclonal population of T cells comprising the T cell of claim 40.
- 35 45. A pharmaceutical composition comprising the T cell of claim 40 and a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.

46. An isolated protein molecule comprising the binding domain of a T cell receptor specific for an MHC-peptide complex, the complex comprising the epitope of claim 1.
47. The protein of claim 46, wherein the protein is multivalent.
48. An isolated nucleic acid encoding the protein of claim 46.
- 5 49. A recombinant construct comprising the nucleic acid of claim 48.
50. A host cell expressing the recombinant construct of claim 33 or 49.
51. The host cell of claim 50, wherein the host cell is a dendritic cell, macrophage, tumor cell, or tumor-derived cell.
52. The host cell of claim 50, wherein the host cell is a bacterium, fungus, or 10 protozoan.
53. A pharmaceutical composition comprising the host cell of claim 50 and a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.
54. A vaccine or immunotherapeutic composition comprising at least one component selected from the group consisting of the epitope of claim 1; the composition of claim 18, 32, 45, or 15 53; the construct of claim 33; the T cell of claim 40, and the host cell of claim 50.
55. A method of treating an animal, comprising:  
administering to an animal the vaccine or immunotherapeutic composition of claim 54.
56. The method of claim 55, wherein the administering step comprises a mode of delivery selected from the group consisting of transdermal, intranodal, perinodal, oral, intravenous, 20 intradermal, intramuscular, intraperitoneal, mucosal, aerosol inhalation, and instillation.
57. The method of claim 55, further comprising a step of assaying to determine a characteristic indicative of a state of a target cell or target cells.
58. The method of claim 57, comprising a first assaying step and a second assaying 25 step, wherein the first assaying step precedes the administering step, and wherein the second assaying step follows the administering step.
59. The method of claim 58, further comprising a step of comparing the characteristic determined in the first assaying step with the characteristic determined in the second assaying step to obtain a result.
60. The method of claim 59, wherein the result is selected from the group consisting of: evidence of an immune response, a diminution in number of target cells, a loss of mass or size 30 of a tumor comprising target cells, a decrease in number or concentration of an intracellular parasite infecting target cells.
61. A method of evaluating immunogenicity of a vaccine or immunotherapeutic 35 composition, comprising:

administering to an animal the vaccine or immunotherapeutic composition of claim 54; and evaluating immunogenicity based on a characteristic of the animal.

62. The method of claim 61, wherein the animal is HLA-transgenic.

5 63. A method of evaluating immunogenicity, comprising: *in vitro* stimulation of a T cell with the vaccine or immunotherapeutic composition of claim 54; and evaluating immunogenicity based on a characteristic of the T cell.

64. The method of claim 63, wherein the stimulation is a primary stimulation.

10 65. A method of making a passive/adoptive immunotherapeutic, comprising: combining the T cell of claim 40 or the host cell of claim 50 with a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.

66. A method of determining specific T cell frequency comprising the step of contacting T cells with a MHC-peptide complex comprising the epitope of claim 1.

15 67. The method of claim 66, wherein the contacting step comprises at least one feature selected from the group consisting of immunization, restimulation, detection, and enumeration.

68. The method of Claim 66, further comprising ELISPOT analysis, limiting dilution analysis, flow cytometry, *in situ* hybridization, the polymerase chain reaction or any combination thereof.

20 69. A method of evaluating immunologic response, comprising the method of claim 66 carried out prior to and subsequent to an immunization step.

70. A method of evaluating immunologic response, comprising: determining frequency, cytokine production, or cytolytic activity of T cells, prior to and subsequent to a step of stimulation with MHC-peptide complexes comprising the epitope of claim 1.

25 71. A method of diagnosing a disease comprising: contacting a subject tissue with at least one component selected from the group consisting of the T cell of claim 40, the host cell of claim 50, the antibody of claim 36, the protein of claim 46; and diagnosing the disease based on a characteristic of the tissue or of the component.

72. The method of claim 71, wherein the contacting step takes place *in vivo*.

30 73. The method of claim 71, wherein the contacting step takes place *in vitro*.

74. A method of making a vaccine, comprising: combining at least one component selected from the group consisting of the epitope of claim 1; the composition of claim 18, 32, 45, or 53; the construct of claim 33;

the T cell of claim 40, and the host cell of claim 50, with a pharmaceutically acceptable adjuvant, carrier, diluent, or excipient.

75. A computer readable medium having recorded thereon the sequence of any one of SEQ ID NOS: X -Y, in a machine having a hardware or software that calculates the physical, 5 biochemical, immunologic, or molecular genetic properties of a molecule embodying said sequence.

76. A method of treating an animal comprising combining the method of claim 55 combined with at least one mode of treatment selected from the group of radiation therapy, chemotherapy, biochemotherapy, and surgery.

10 77. An isolated polypeptide comprising an epitope cluster from a target-associated antigen having the sequence as disclosed in Tables 25-44, wherein the amino acid sequence consists of not more than about 80% of the amino acid sequence of the antigen.

78. A vaccine or immunotherapeutic product comprising the polypeptide of claim 78.

79. An isolated polynucleotide encoding the polypeptide of claim 78.

15 80. A vaccine or immunotherapeutic product comprising the polynucleotide of claim 80.

81. The polynucleotide of claim 79 or 80, wherein the polynucleotide is DNA.

82. The polynucleotide of claim 79 or 80, wherein the polynucleotide is RNA.

FIG. 1A

1	CTAG_HUMAN	NY-ESO	50
	AAD05202 - CAG-3	(1) MQAERGCTGGTCAEPFCCGTCGTCGCGAGGCGGAA	
	CAA11044 - LAGE-1a	(1) MQAEGCTGGTCAEPFCCGTCGTCGCGAGGCGGAA	
	CAA110194 - LAGE-1s	(1) MQAEGCTGGTCAEPFCCGTCGTCGCGAGGCGGAA	
	CAA11043 - LAGE-1b	(1) MQAEGCTGGTCAEPFCCGTCGTCGCGAGGCGGAA	
	CAA110196 - LAGE-1L	(1) MQAEGCTGGTCAEPFCCGTCGTCGCGAGGCGGAA	
	AH02883 CT-2	(1) MQAERGCTGGTGGSTGDAAGPGGGPIDPGGGNAGGPGGEAGATGGGPRGAGA	100
	Consensus	(1) MQAEGRGTTGGSTGDAAGPGGGPIDPGGGNAGGPGGEAGATGGGPRGAGA	
51			
			51
	CTAG_HUMAN	NY-ESO	
	AAD05202 - CAG-3	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	CAA11044 - LAGE-1a	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	CAA110194 - LAGE-1s	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	CAA11043 - LAGE-1b	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	CAA110196 - LAGE-1L	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	AH02883 CT-2	(51) ARAGCCTGAAKREPGGAA\$LNCCREGATGEPSPRF\$F\$AMP\$PM	
	Consensus	(51) ARASGPRGGAPRGPHGGAASAQDGRCPGARRPDSRLLQLHITMPFSSPM	

## FIG. 1B

CTAG_HUMAN	NY-ESO	(101)	FAAFAATSENQDAPPPPEVDSV	101
AAD05202	- CAG-3	(101)	FAAFAATSENQDAPPPPEVDSV	150
CAA11044	- LAGE-1a	(101)	FAAFAATSENQDAPPPPEVDSV	
CAA10194	- LAGE-1s	(101)	FAAFAATSENQDAPPPPEVDSV	
CAA11043	- LAGE-1b	(101)	FAAFAATSENQDAPPPPEVDSV	
CAA10196	- LAGE-1L	(101)	FAAFAATSENQDAPPPPEVDSV	
AAH02833	CT-2	(101)	FAAFAATSENQDAPPPPEVDSV	
Consensus		(101)	EAELVRRILSRDAAPLPRPGAVLKDFTVSGNLLFIRLTAADHRQLQLSIS	
CTAG_HUMAN	NY-ESO	(151)	-----	151
AAD05202	- CAG-3	(151)	-----	200
CAA11044	- LAGE-1a	(151)	-----	
CAA10194	- LAGE-1s	(151)	-----	
CAA11043	- LAGE-1b	(151)	VGWGLGASASPEGQKARDLRTPKHKVSE	
CAA10196	- LAGE-1L	(151)	VGWGLGASASPEGQKARDLRTPKHKVSE	
AAH02833	CT-2	(151)	VGWGLGASASPEGQKARDLRTPKHKVSE	
Consensus		(151)	SCLQQQLSLIMWITQCLFLPVFLAQ PSGQRR-----	

201

CTAG	HUMAN	NY-ESO	(181)	-----
AAD05202	-	CAG-3	(181)	-----
CAA11044	-	LAGE-1a	(181)	-----
CAA10194	-	LAGE-1s	(181)	-----
CAA11043	-	LAGE-1b	(201)	ENVMSAPHI
CAA10196	-	LAGE-1L	(201)	ENVMSAPHI
AAH02833	CT-2		(201)	ENVMSAPHI
Consensus				(201)

FIG. 1C

4/22

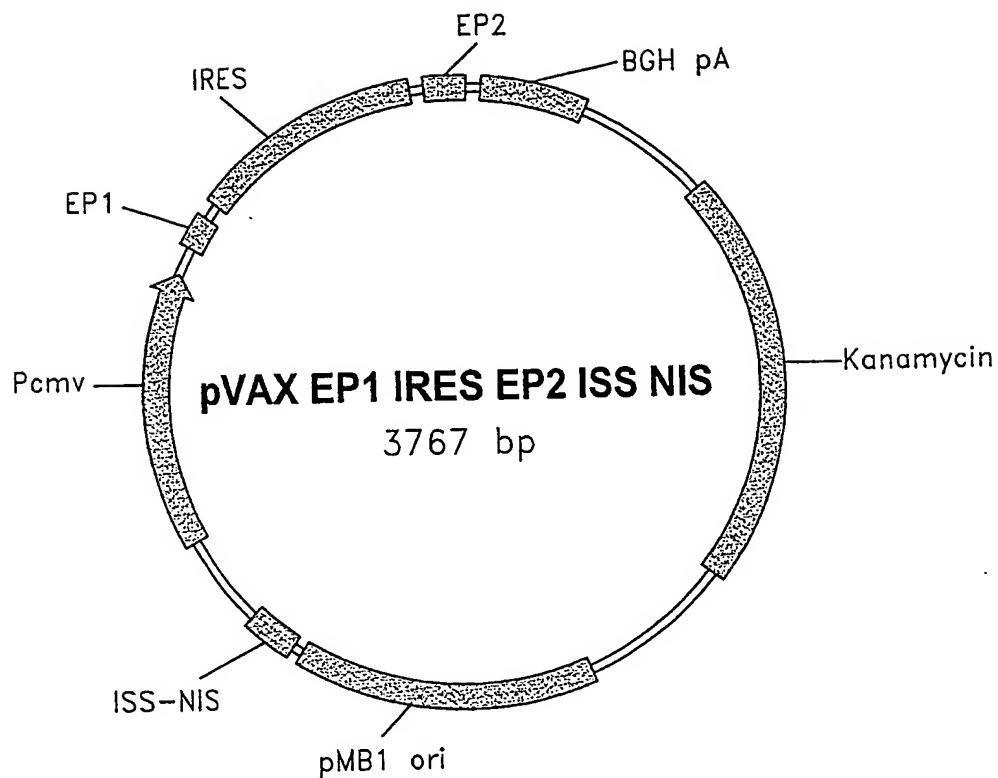
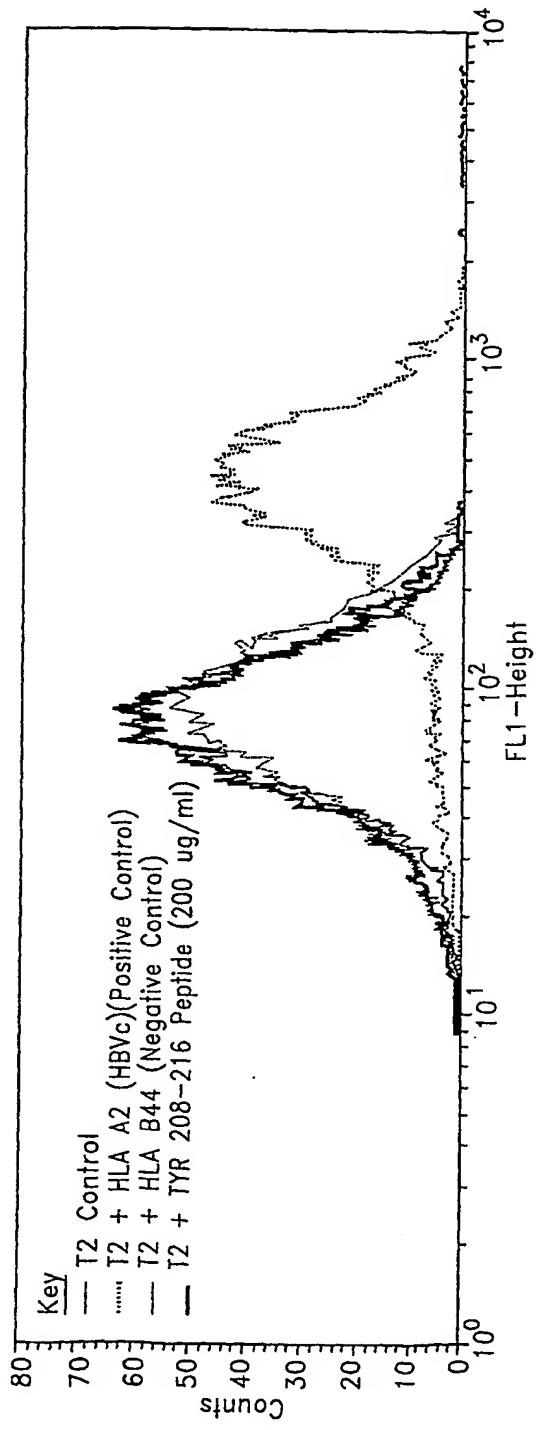


FIG. 2

FIG. 3A

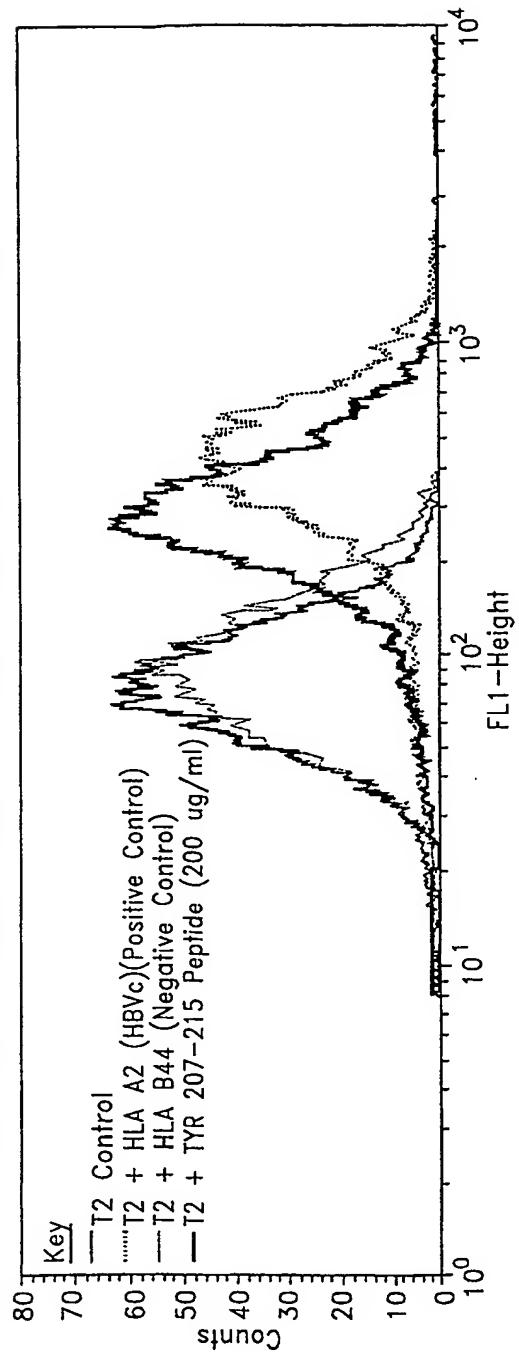
**FACscan Analysis of Binding Assay to Determine the Binding Ability of Tyrosinase 208-216 Peptide to MHC Class 1**



$$\begin{aligned}
 F1 \text{ (HLA A2 Peptide)} &= 3.13 \\
 F1 \text{ (TYR 208-216 Peptide)} &= 0.01
 \end{aligned}$$

FIG. 3B

**FACscan Analysis of Binding Assay to Determine the Binding Ability of Tyrosinase 207-215 Peptide to MHC Class 1**



$$\begin{aligned}
 F1 \text{ (HLA A2 Peptide)} &= 3.13 \\
 F1 \text{ (TYR 207-215 Peptide)} &= 2.00
 \end{aligned}$$

7/22

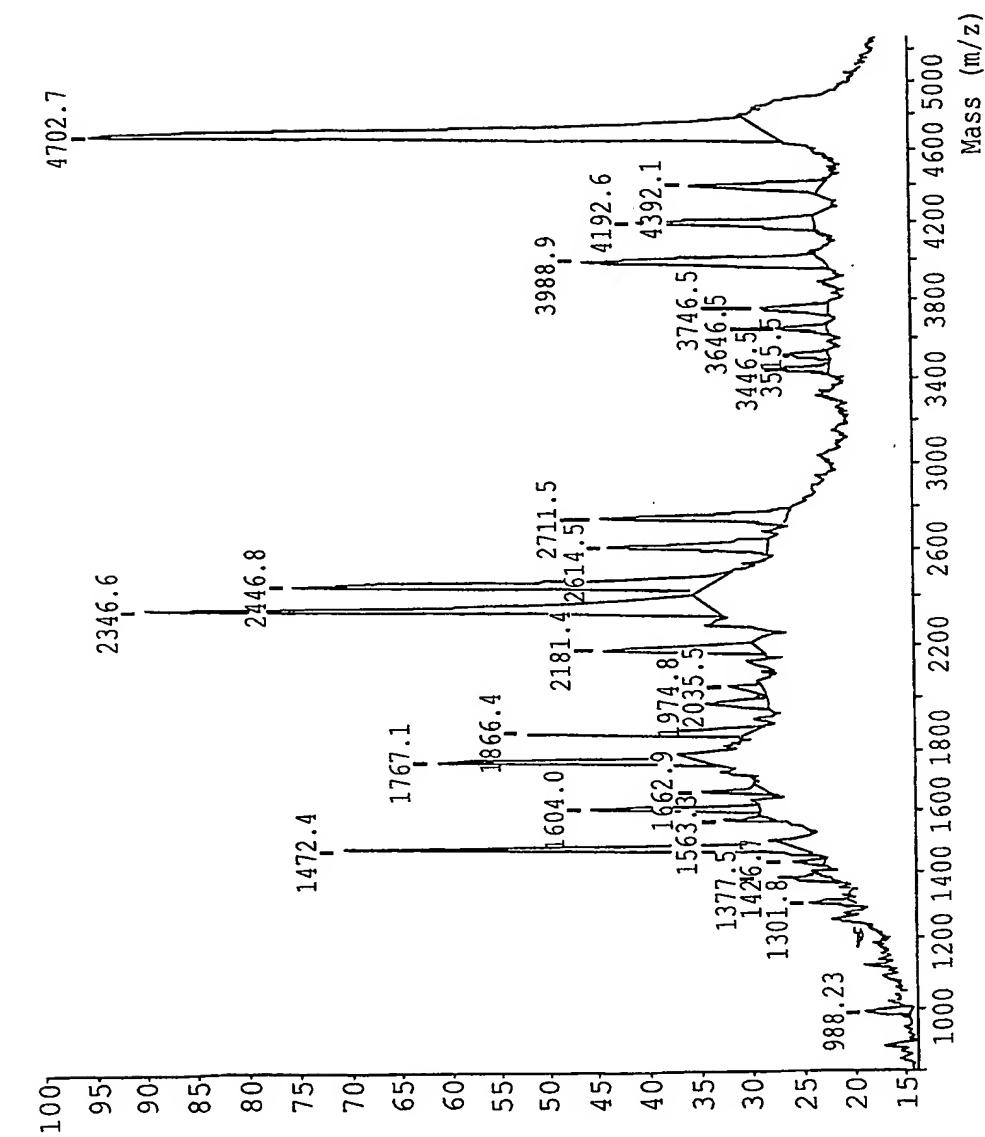
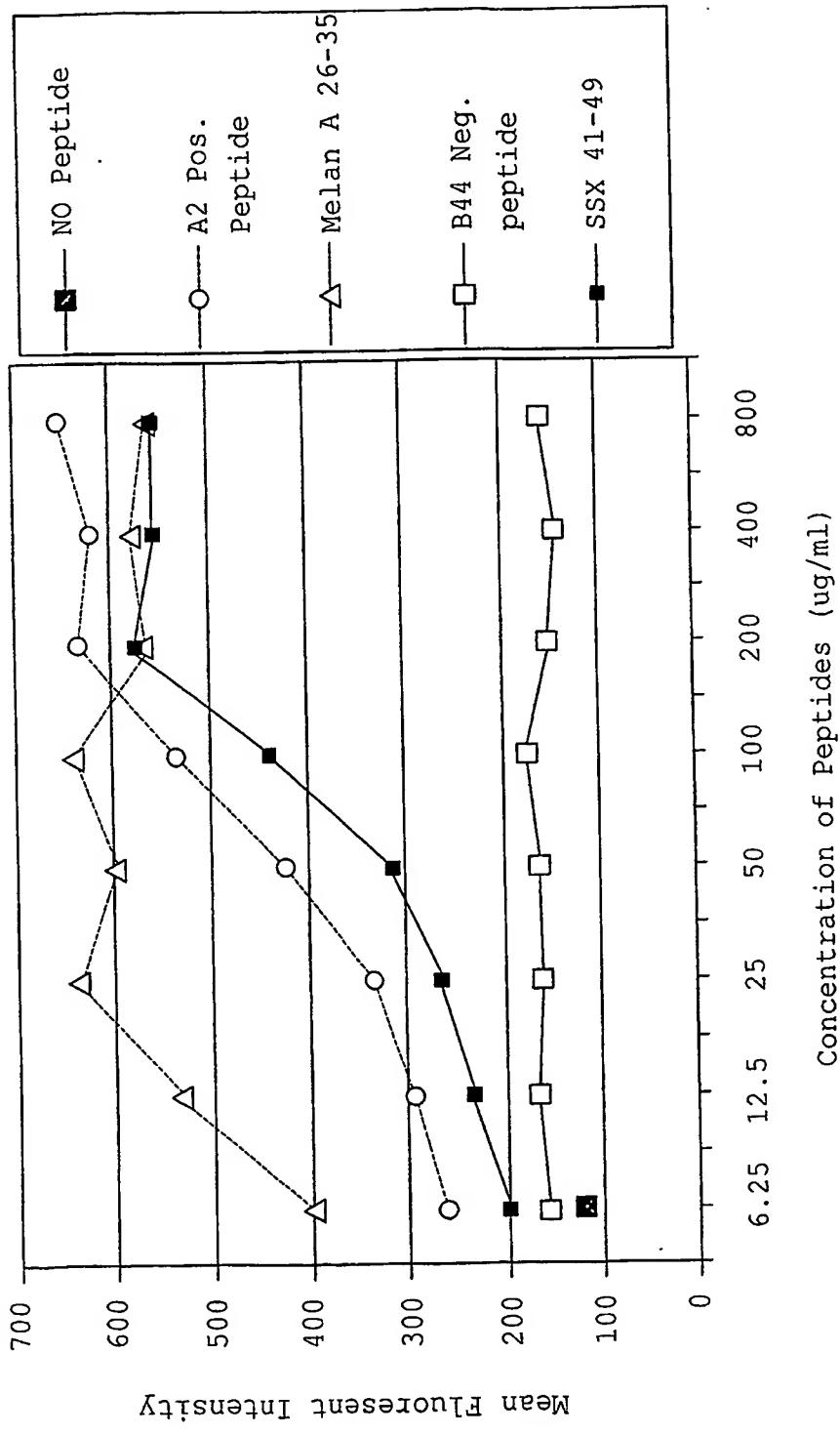


FIG. 4

SUBSTITUTE SHEET (RULE 26)

*FIG. 5*  
Comparison of Peptides Binding Affinity to HLA A2



9/22

FIG. 6

**SSX2<sub>41-49</sub> specific lysis by CTL from peptide injected HD1 mice**

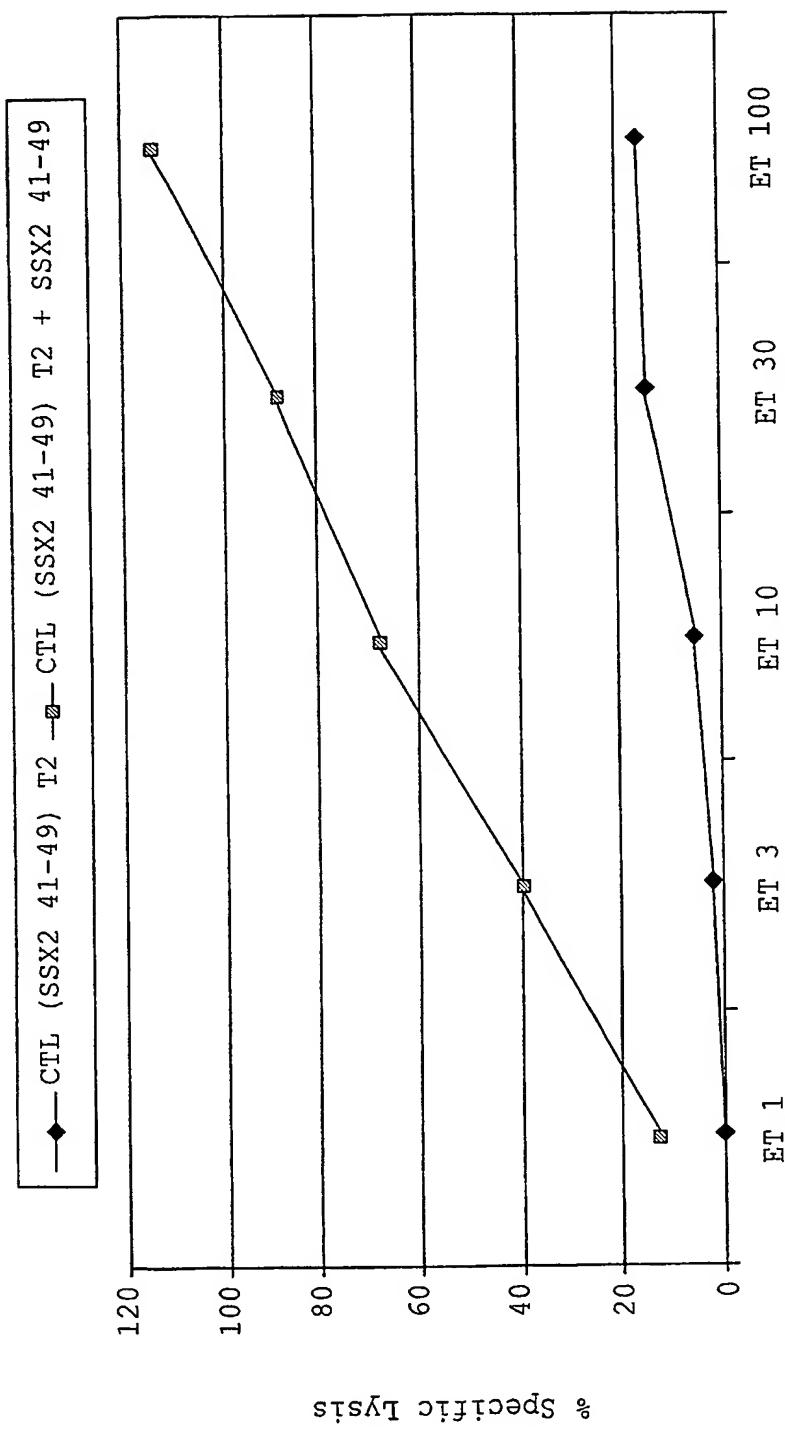
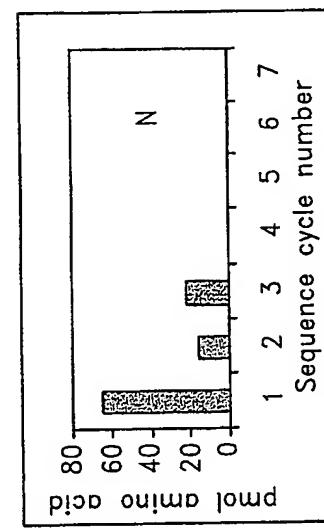
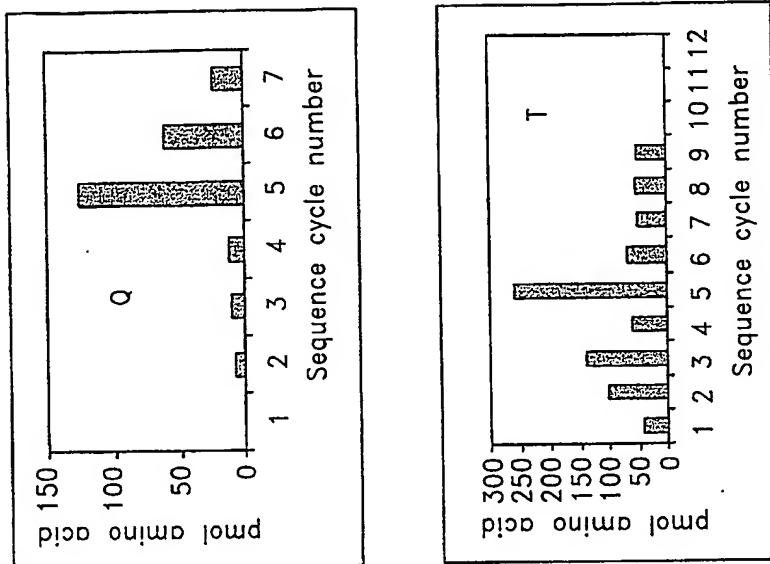
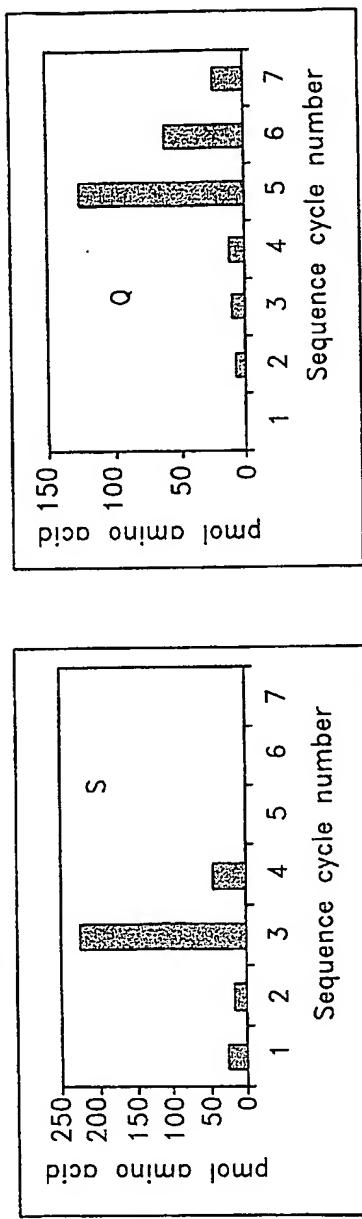
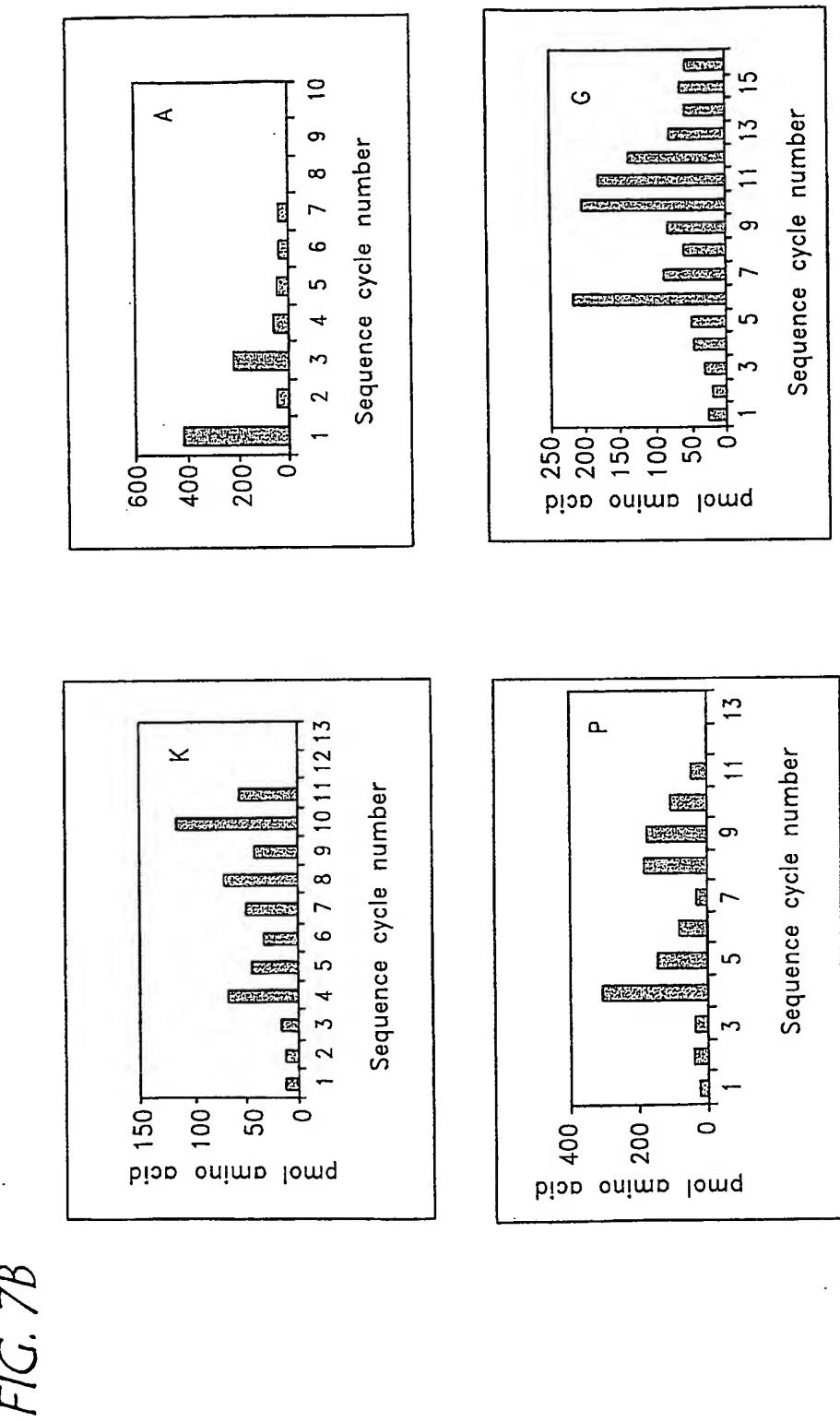


FIG. 7A

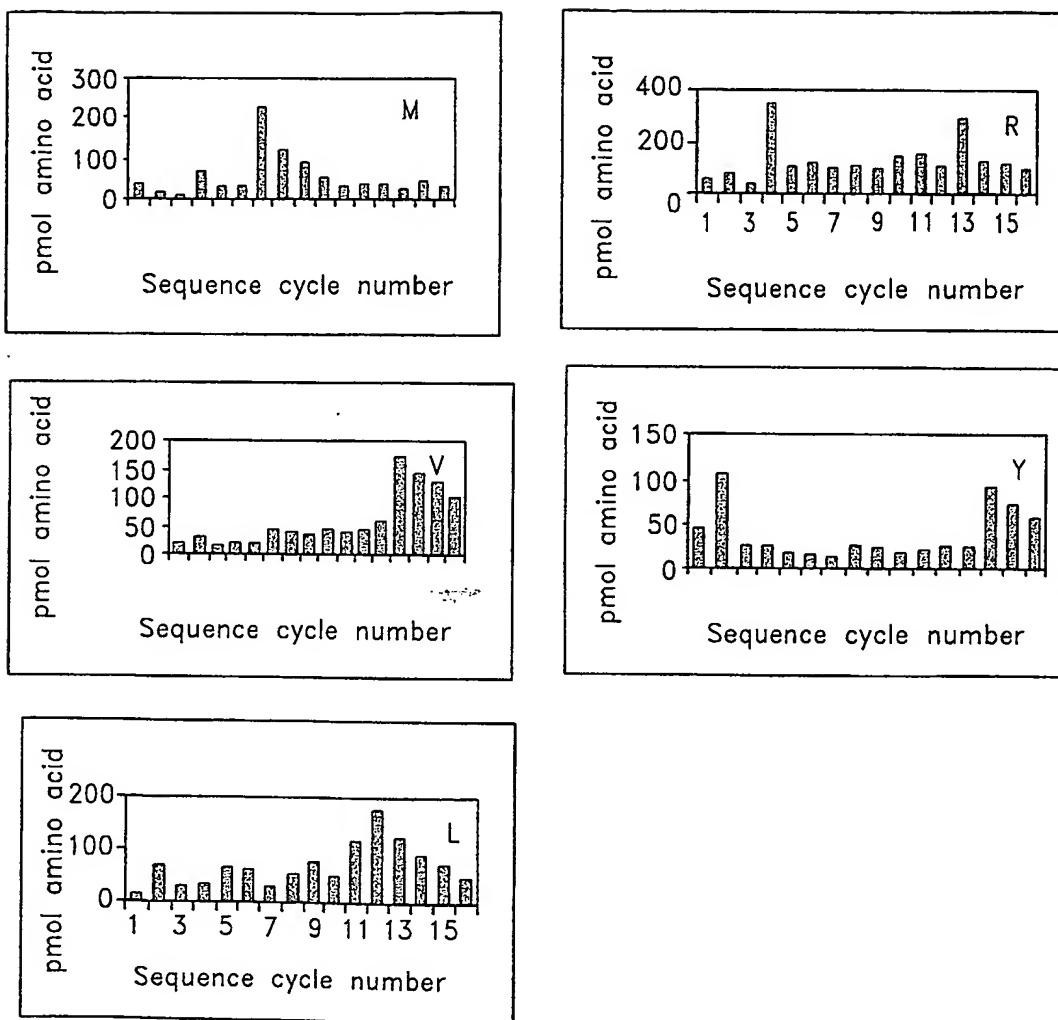
163-AF**S**P**Q**GMPEGDLVYV**N**YARTEDFFKLERDM-192

Pool sequencing of PSMA\_163-192 Digested for 60 min by proteasome



Pool sequencing of PSM A\_163-192. Digested for 60 min by proteasome

12/22

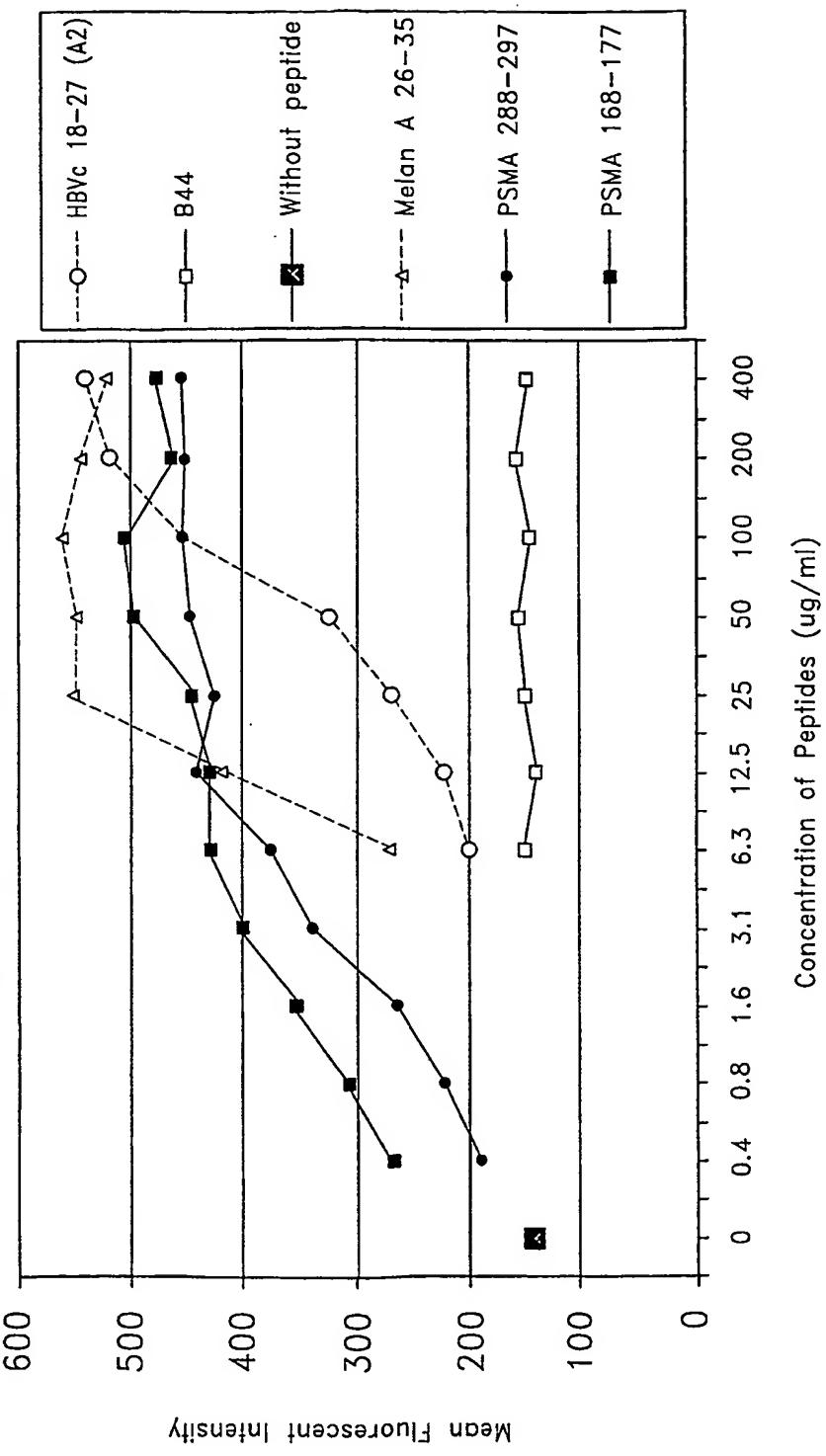


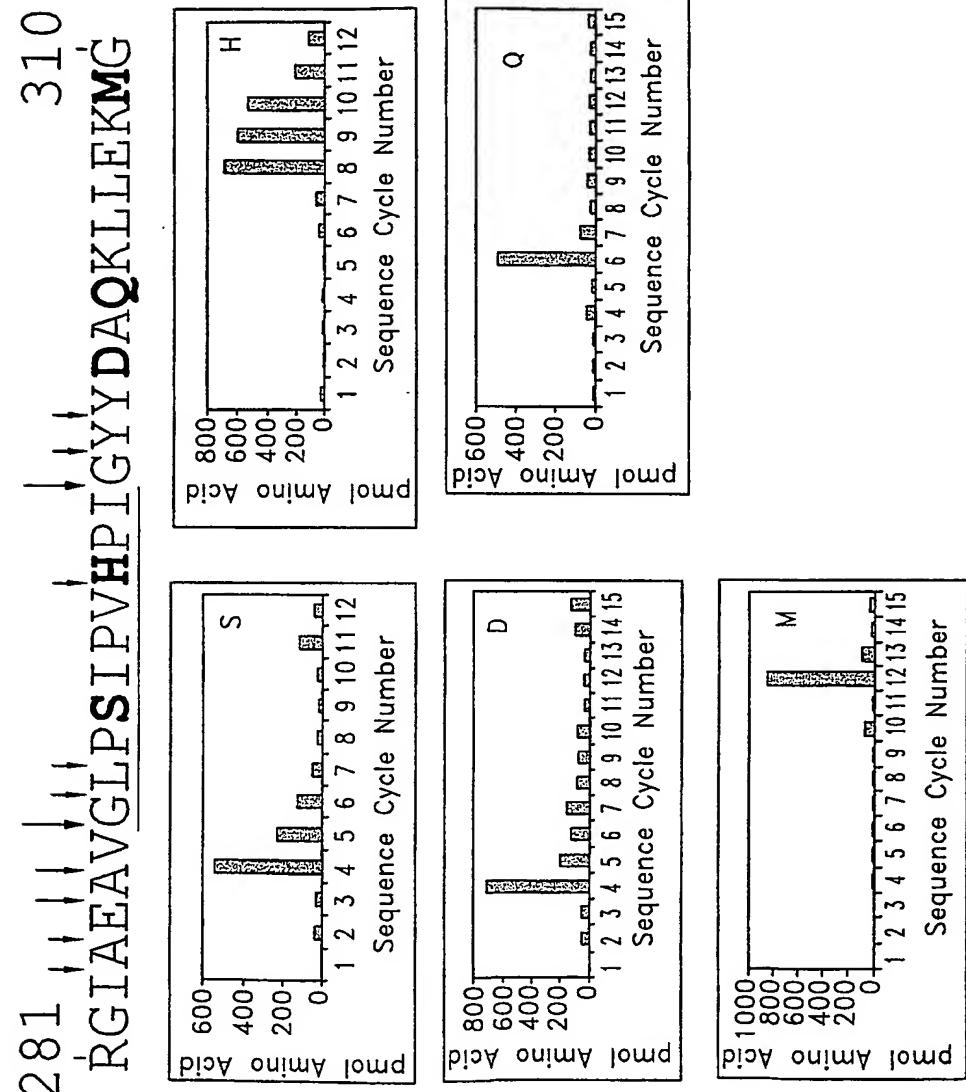
Pool sequencing of PSMA\_163-192 Digested for 60 min by proteasome

FIG. 7C

SUBSTITUTE SHEET (RULE 26)

FIG. 8

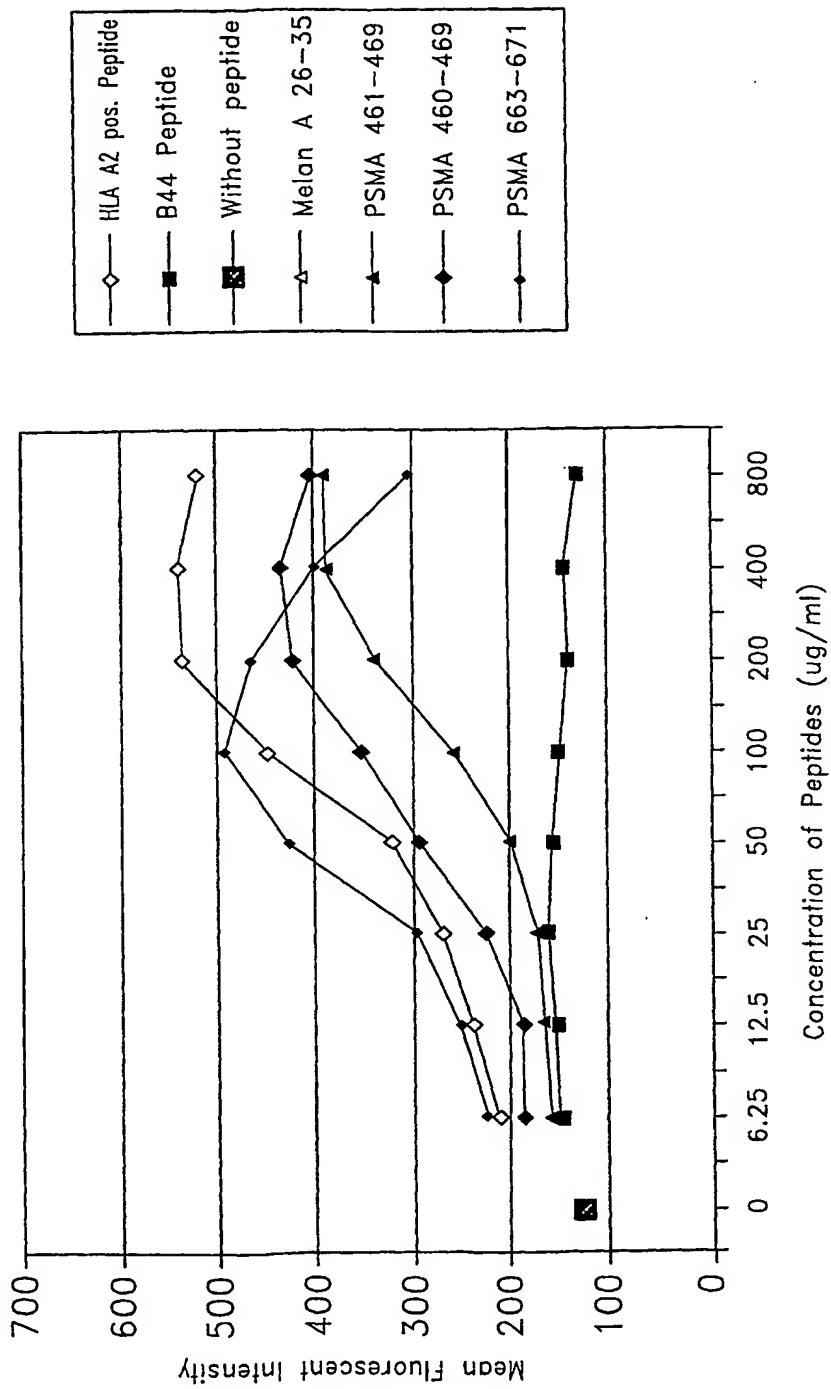




Pool sequencing of PSMA\_281\_310 Digested for 60 min by Proteasome

15/22

*FIG. 10*  
**Comparison of Peptides Binding  
 Affinity to HLA A2  
 by Binding Assay**



Autologous DC Present A1  
Peptide to CD8 T cell

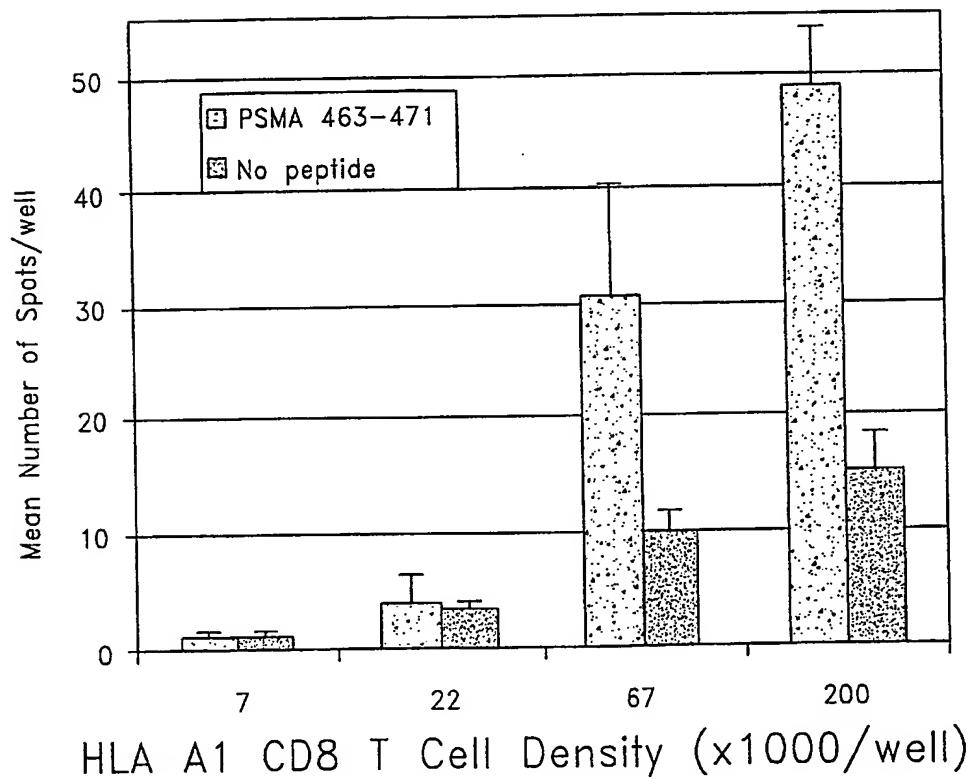


FIG. 11

## Secretion of IFNgama Was Blocked by Anti-A1 Antibody

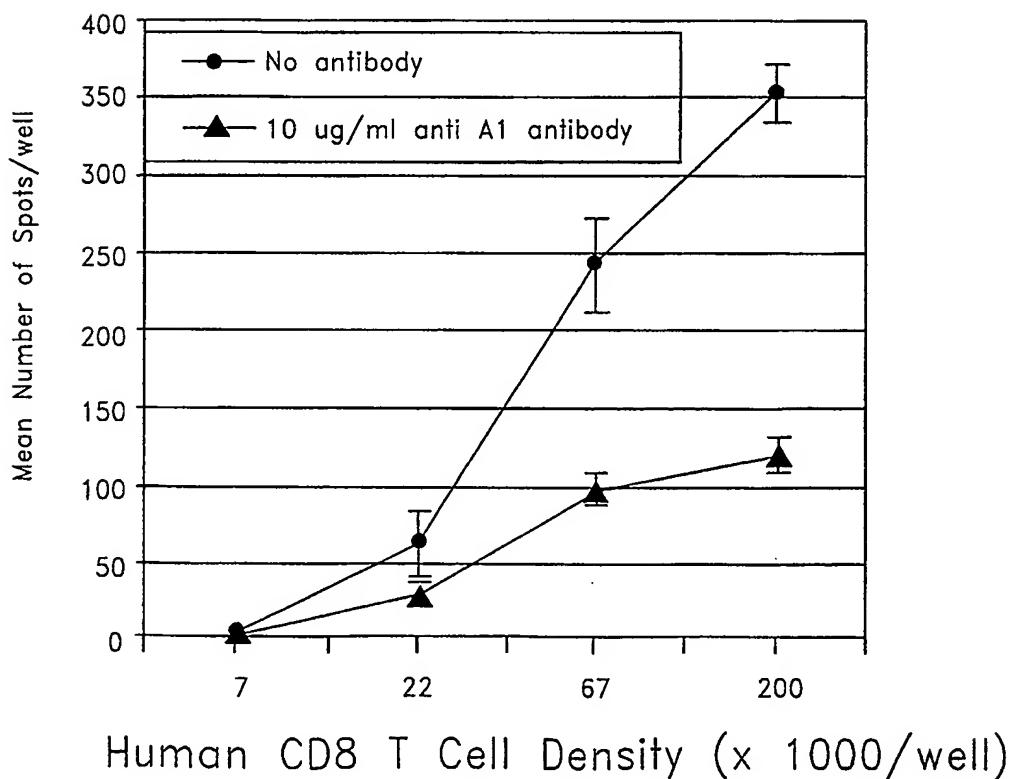


FIG. 12

FIG. 13

Comparison of Peptides Binding Affinity  
to HLA A2 by Binding Assay

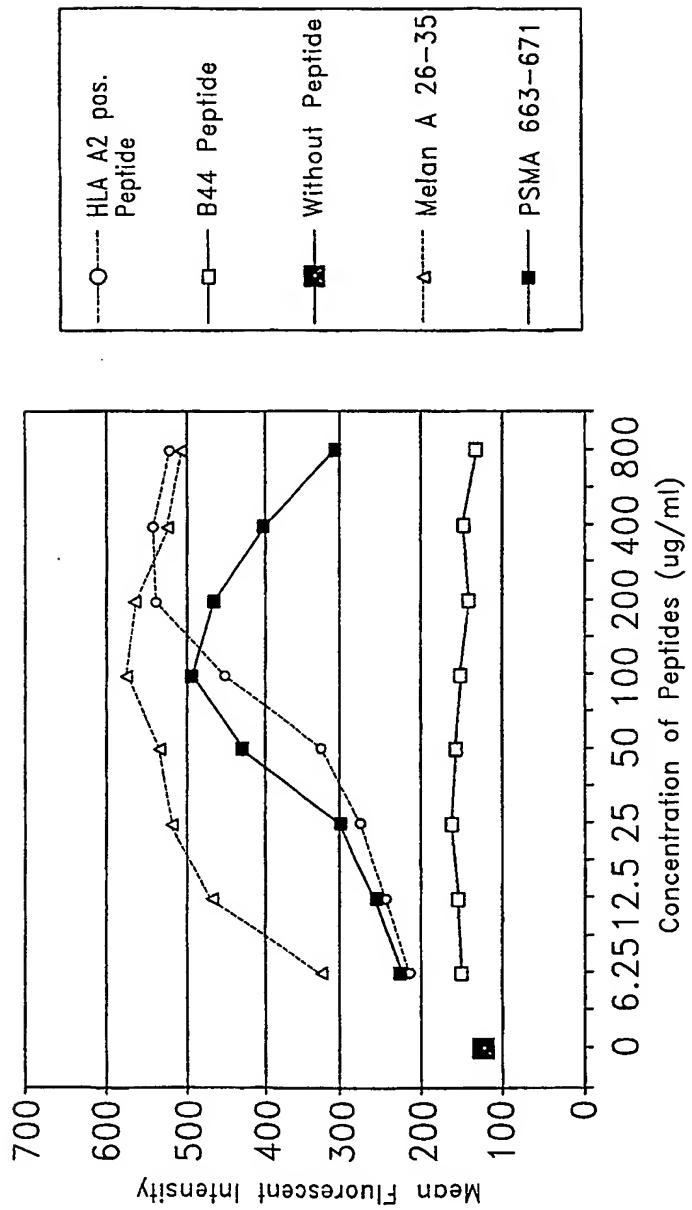
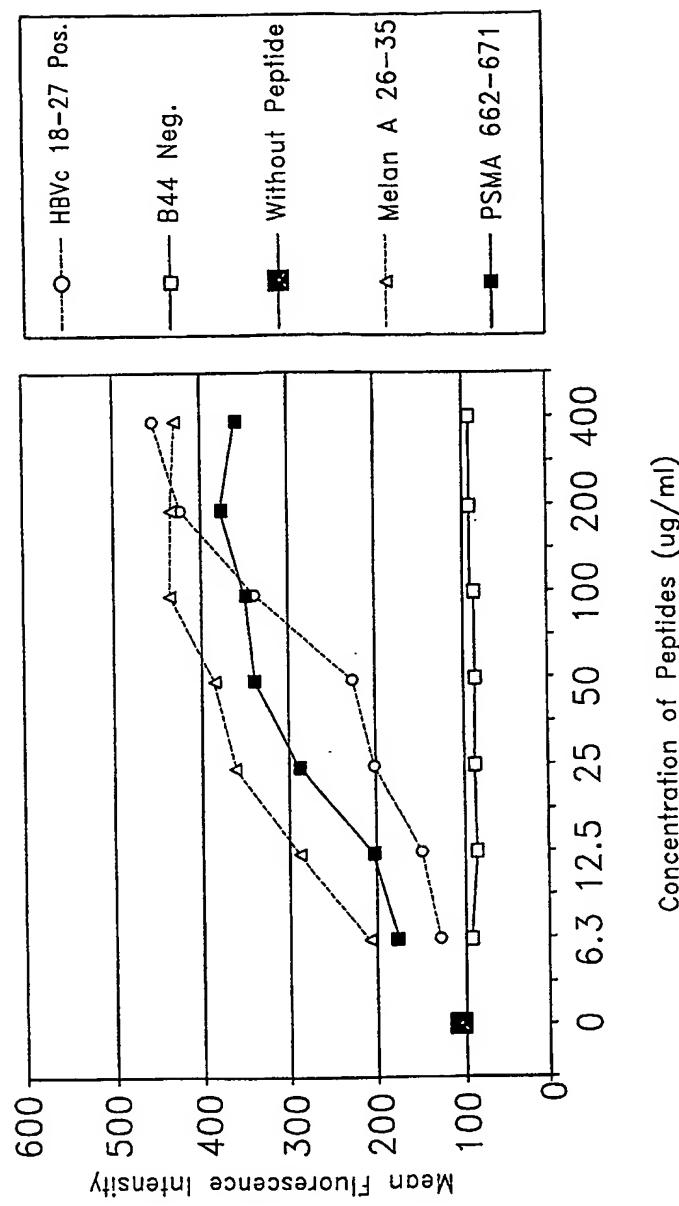
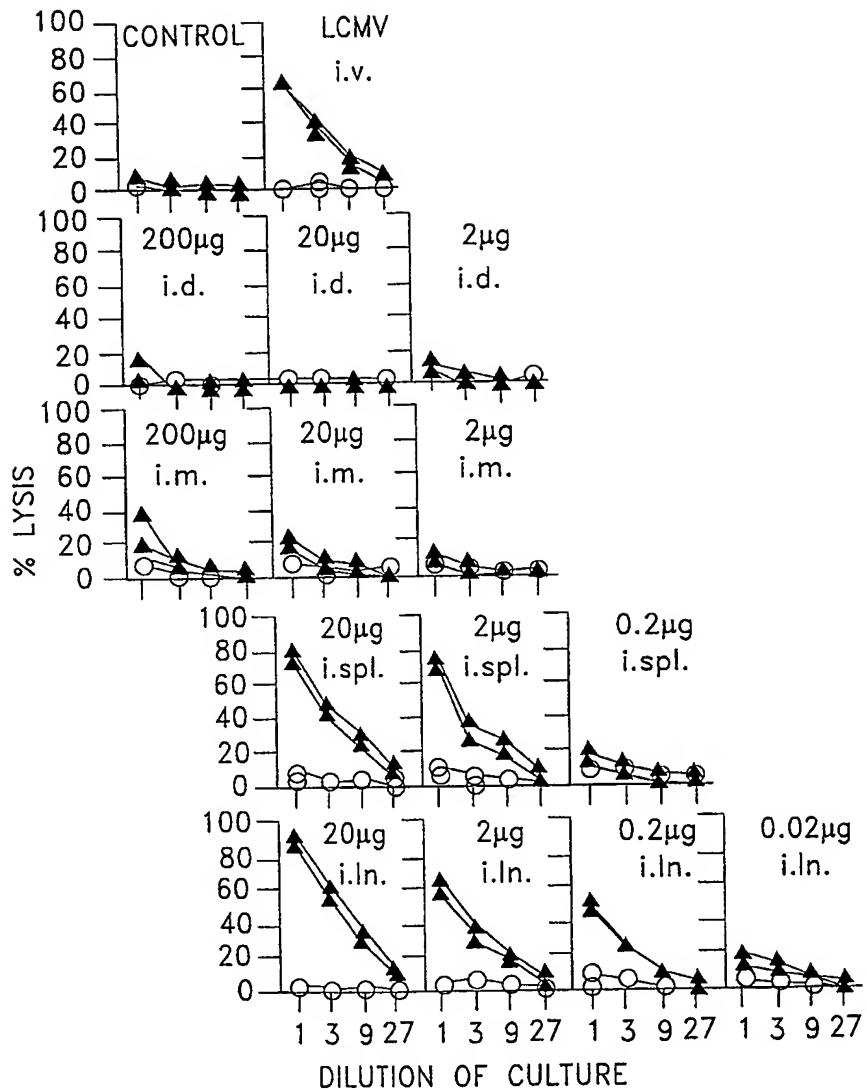


FIG. 14

Comparison of Peptides Binding Affinity  
to HLA A2 by Binding Assay

20/22

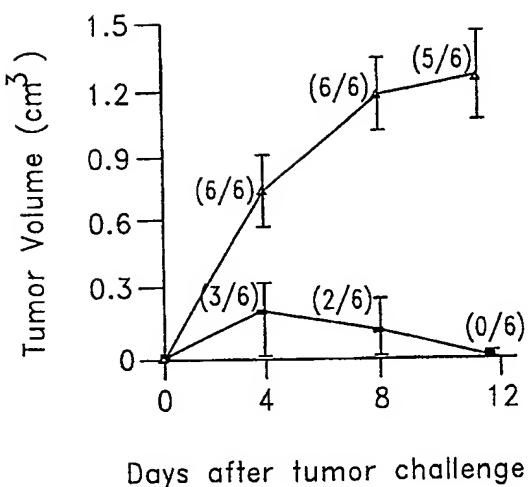


Graphs show lysis of unpulsed EL4 cells (open circles) and EL4 cells pulsed with gp33 peptide (solid triangles). Symbols represent individual mice and one of three similar experiments is shown.

FIG. 15

SUBSTITUTE SHEET (RULE 26)

21/22



Mean tumor volumes  $\pm 1\text{SD}$  are shown for mice immunized with pEFGPL33A DNA (solid circles) or control pEGFP-N3 DNA (open triangles). Numbers in brackets indicate number of mice with tumors/total number of mice in group. One of two similar experiments is shown.

*FIG. 16*

22/22

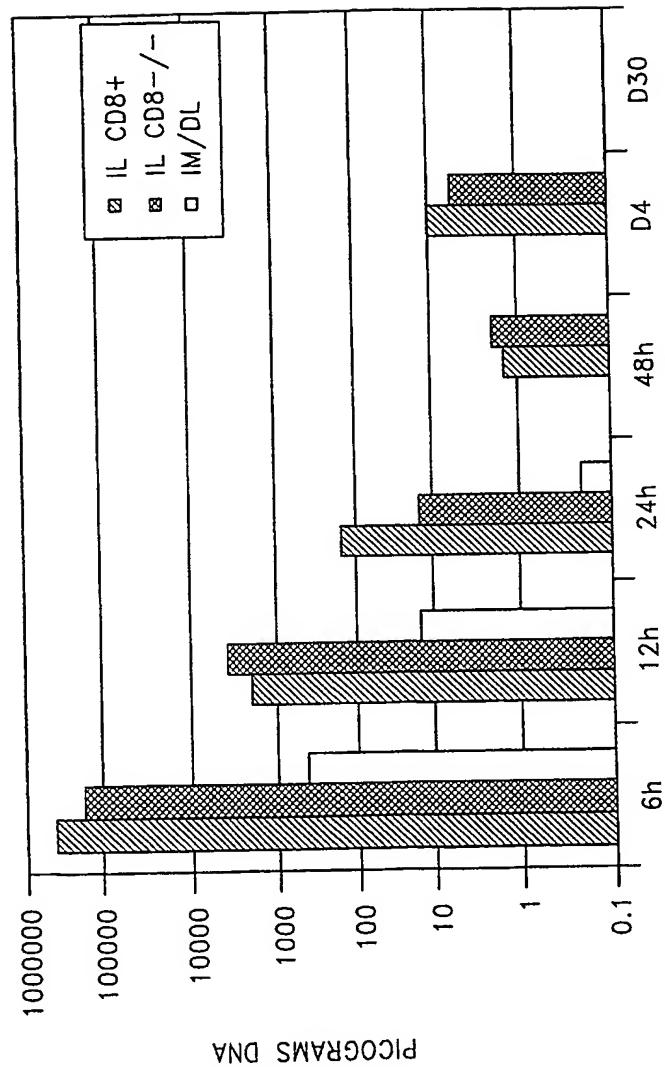


FIG. 17

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**